

Query Match  
Best Local Similarity

	Matches	1539;	Conservative	0;	Mismatches	271;	Indels	6;	Gaps	2;
OY	127	AGGACAGCTGGCCTGAAGCTCAGAGCGCGGGCGCTGCGCCATGAGCCGCCACACTGGGCTGTC	186							
DB	61	AGGGCTGTGGACTGAAGATTAAACCCCTGGGTGTCTGCCATGGCCCCACACTGGGCTGTGC	120							
OY	187	TGGCTCTGGCACCAAGCGCTGGGGGCGCTGGGGCATTTGGGGCTGAGGTGTGGGAACCTT	246							
DB	121	TGGCTCTGGCACCAAGCGCTGGGGGCGCTGGGGCATTTGGGGCTGAGGTGTGGGAACCTT	180							
OY	247	GTCCGGGTAAAGACAGTGTCTTTGGGGGAGCTGGCCACGGTACTAGAGCGGTCTCCAG	306							
DB	181	GTGCCCGCGAAGACAGATCTTTCTGGGGAGCTGTCAAGTACTGAGGGGGTCTCCAG	240							
OY	307	ACCGGACATCCAGAGACTTCTGACACTGACGCTGACGGAGCCCACTGGGCTTCTGACGTG	366							
DB	241	ACAGGACATCCAGAGACTTCTGACACTGACGCTGACAGAACATCTTGGCTTTATATGTG	300							
OY	367	GGGCGCCGAGAGGCGCCCTGTGTGCTTTCAGATGAGAGGCCCTGGAGGCTGCAAGGAGGATC	426							
DB	301	GGGCGCCGAGAGGCGCGCTGTGTGCTTTCAGATGAGAGGCTTGGAGCTGCAAGAGCGATC	360							
OY	427	TCCGAGGAGGCGCCCGCGGAGAAAGACAGTGTATCCAGAAAAGGAAACACAG	486							
DB	361	TCTTGGAGAGGCTCCAGCTGAGAAAGAAATTTGAATGTACCCGAAAGGAAAGACAAACAG	420							
OY	487	ACGAGAGCTTCAACTTCATCCGCTTCTGACAGCCCTACAAAGCTCCACCTGTATGTC	546							
DB	421	ACGGAATGCTTCAACTTCATCCGCTTCTGACAGCCATACAAATCTCTCCATCTGTATGTC	480							
OY	547	TGTGGCACTTACGCTTTCAGCCCAAGTGCACCTACGTCAACATNGTCCACTTCATCTTG	606							
DB	481	TGCGGCACCTTATGCTTCCACGCCCAAGTGCACCTACATCAACAGTCCACGTCACCTTG	540							
OY	607	GAGCATGAGAGTTTGAAGATGGGAGGGGCAAGTGTCCCTATGACCCAGCTAAGGGCCAT	666							
DB	541	GACCGTCAAGAAATTTGAGATGGGAGGGGAAAGGTAAATGCCATATGACCCAGCTAAGGGTAC	600							
OY	667	GCTGGCCTTCTTGTGATGTGTAGCTGTACTGTGGCCACACTCAACAACTTCTGGGACG	726							
DB	601	ACCGGACTCCTTGTGAGAGGTGTAGCTGTACTGACACACACTCAATTAATCTCTGGGACA	660							
OY	727	GAACCCATTTATCTGCGCTAATATGGGGGCGCCACACTCCATGSAAGACAGATACCTGGCC	786							
DB	661	GAGCGGTTATCTTGCATACATGGGGGCGCCACACTCCATCAAGACAGATACCTGGCT	720							
OY	787	TTTTGGCTCAAGAACCTCACTTGTATAGCTCTGCGCTATGTAAGCTGAGAGTGGGACG	846							
DB	721	TTTTGGCTCAAGAACCCCACTTTGTAGCTGTGCTTGTCCCGAAGGTGGGAAGC	780							
OY	847	TTTACGAGGGGAGCAGACAGAGTCTACTTCTTTCAGGAGAGCGGGCAGTGGATCCGAC	906							
DB	781	TTTACGAGGGGAGCAGATGACAGATCTACTTCTTTCAGTGAAGCGGGCAGTGGAGTATAC	840							
OY	907	TGCTATGCCGAGACAGTGTGGCTGTGTGTGGCCCTGTCTCAAGGGGGAATATGGGGGCG	966							
DB	841	TGCTATTTCCGAGCAGGTGTGGCTGTGTGTGGGAAGTCTGTAAAGGGTACATGGGGGA	900							
OY	967	GCAGGAGCCCTCAGAGAGAGTGGACACAGTTCCTGAAGGGCGGGCTGGATGCTCTGCC	1026							
DB	901	GCAGGAGCGCTCAGAGAAATATGACAGAGCTTCTCGAAAGGCTCGGTTGGTGTGCTCACGC	960							
OY	1027	CCGAAGTGGCAGCTTACTTCAACAGCTGACAGGGGATGACACACCCTGCAGGACACTCC	1086							
DB	961	CCTGAAGTGGAGGCTTACTTCAACAGCTGAAAGGGGATGACACACCCTGCAGGAGCCTCT	1020							
OY	1087	TGGCACAAACACACTTCTTTGGGGTTTTTCAAGCACAGTGGGGTGAATGTACTGTGCG	1144							
DB	1021	TGGCACAAACACACTTCTTGGGGTTTTTCAAGGGCGAATGGGGGAATATGACCTGTGCT	1080							
OY	1147	GCATCTGTGATTACAGTGTGAAAGATATCCACGCGGTTGTTGAGGGCGCCCTATTAAGAG	1206							
DB	1081	GCATTTGTGATTAACAGTTTGGAAACAGATCCAGCAAGTGTTTGAGGGTCCCTTCAAGAGAG	1144							

**THIS PAGE BLANK (USPTO)**

Wed Aug 20 09:19:27 2003

us-10-002-05

1207 TACCATGAGAGAGCCAGAGTGGAGCCGCTACAGTACCCTGTACCAGCCCTCGGCT 1266  
1141 TACAGTAGAGAGCCAGAGTGGAGCCGCTATACAGCCGCTACAGCCGCTCGGCT 1200  
1267 GGCTGCTGATTAACAACCTGGCCAGCCGCTACAGCAGCTCCCTGGAGCTACCC 1326  
1201 GGTTGCTGATCAACAACCTGGCCAGAGACAGTACAGCAGTCCCTGGAGCTGCG 1260  
1327 GACACATCCCTCAACTGCTGAAGAGACCCGCTGATGAGAGAGAGTGGGCTCGG 1386  
1261 GACACACACCCCTCAACTGCTGAAGAGACCCGCTGATGAGAGAGAGTGAAGCTCGG 1320  
1387 TGAAGCCGCCCCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446  
1321 TTGGGCCCCCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1447 GTTACAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506  
1381 GTCCAGAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
1507 CTGCTCAAGGCTGAG 1566  
1441 CTGCTCAAGGCTGAG 1500  
1567 GACAG 1626  
1501 GACAG 1560  
1627 TCCGCTGCTGAG 1686  
1561 TCCGCTGCTGAG 1620  
1687 GACTGTCTCTGAG 1746  
1621 GACTGTCTCTGAG 1680  
1747 G---CCGTTGGTGGCCACTTGGATCTTACTGATCCAGATGATGATGATGATGAT 1803  
1681 GCCACACACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1740  
1804 TCAGGCAATTGCAACCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1863  
1741 TCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1797  
1864 ACAGTGGTGGGAG 1923  
1798 ACCGTTGCTGAG 1857  
1924 CCCGACTGCAACCCCG 1939  
1858 GCCCACTGAGCTTGC 1873

**THIS PAGE BLANK (USPTO)**





JOURNAL Patent: WO 0078802-A 13 28-DEC-2000;  
Curagen Corporation (US)  
FEATURES Location/Qualifiers  
source 1..2284  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
166..1956  
CDS  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC24872.1"  
/db\_xref="GI:12405795"  
/translation="MAPHRYVLLAARLMLGICAEVWNNVLPKRTVSSGELATYVR  
FSQIGIDFLITLLEPTGLLVGAREALFASMEALELQGAISWEAPVEKTEKIQK  
GKNQTECFNFIREFLOPYNASHLYVCTYAFQPCITYVMILFTELEHFEEDKQKCP  
YDPAGHAGLIVDELVSATLNNLSTGEPIILRNMGPHNSMKEYLEALNEPDKVPS  
AYPESVSFTDDDKVYEFPERAVESPCYEOVVAVARAOKDMGAPRILOKMT  
TEIKRLACSAENMOLYFNOLQAMHTIDTSMHTNPFVGVROKMGDYLSICRYOL  
EELQVFPKYEYHEBAOKNDRTDPIPSRPGSCINNMHRHGTSTSLLEPDLIN  
FVKIPLMEBOVGPWRSPRLVKKGTNTHLADVTGLDQATYVLELIGTDGWLK  
AVSLGPMVHLIEQLFDQEPKRSVLISQSKLFPAGSRSLVOLPVADCIKRYSCAD  
CVLADPYCAMSVTNSRCVAVGSHGSLIIQHVMTSDISGLNLSKSIQSGPLPKNI  
TVVAGTDLVLPCHLSNLALPDSNPESSV"

BASE COUNT 443 a 683 c 689 g 465 t 4 others  
ORIGIN

Query Match 99.8%; Score 2280; DB 6; Length 2284;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCCCTCTACACTCCGCTGCTGATGTGGAACGGGGTTCGGGCTTCGACGGCT 60  
DB 1 CGGCCCTCTACACTCCGCTGCTGATGTGGAACGGGGTTCGGGCTTCGACGGCT 60  
QY 61 ATTGTCTGGCTGGGGAAGGCGACAGCGCGGACCTCCGCTCGACCGCGCGC 120  
DB 61 ATTGTCTGGCTGGGGAAGGCGACAGCGCGGACCTCCGCTCGACCGCGCGC 120  
QY 121 ACCACGAGACAGCTGGCTGAAGCTCAGACCGGGGCGCTGGCCATGGCCACACTGG 180  
DB 121 ACCACGAGACAGCTGGCTGAAGCTCAGACCGGGGCGCTGGCCATGGCCACACTGG 180  
QY 181 GCTGTCTGGCTGTGGCAGCAGGCTGTGGGGCTGGGCTAGGGGCTAGGTGTGGTGG 240  
DB 181 GCTGTCTGGCTGTGGCAGCAGGCTGTGGGGCTGGGCTAGGGGCTAGGTGTGGTGG 240  
QY 241 AACCTTGTGCGCCGTAAGACAGTGTCTTGGGAGCTGGCCACGGTATGACGGCGTTC 300  
DB 241 AACCTTGTGCGCCGTAAGACAGTGTCTTGGGAGCTGGCCACGGTATGACGGCGTTC 300  
QY 301 TCCGACAGCCGGCATCCAGACACTTCTGTACACTGACGCTGACGACCGCCACTGGCGTTCTG 360  
DB 301 TCCGACAGCCGGCATCCAGACACTTCTGTACACTGACGCTGACGACCGCCACTGGCGTTCTG 360  
QY 361 TACGTGGGGGCCCGAGAGGCGCTGTGGCTCAGCATGGAGGCGCTGGAGTGGCAAGA 420  
DB 361 TACGTGGGGGCCCGAGAGGCGCTGTGGCTCAGCATGGAGGCGCTGGAGTGGCAAGA 420  
QY 421 GCGATCTCTGGAGAGGCGCGCTGGAGAGAACTGAGTGTATTCAGAAAGGAGAAAC 480  
DB 421 GCGATCTCTGGAGAGGCGCGCTGGAGAGAACTGAGTGTATTCAGAAAGGAGAAAC 480  
QY 481 AACGACAGCCAGTCTTCAACTTCATCCGCTTCTGTGACGCGCTTACAATGCTCCACCTG 540  
DB 481 AACGACAGCCAGTCTTCAACTTCATCCGCTTCTGTGACGCGCTTACAATGCTCCACCTG 540  
QY 541 TACGCTGTGGACCTAGGCTTTCAGGCCAAGTCACTAGTCAACTGTCTACCTTC 600  
DB 541 TACGCTGTGGACCTAGGCTTTCAGGCCAAGTCACTAGTCAACTGTCTACCTTC 600  
QY 601 ACTTTGAGACATGAGAGTTTGAAGATGGAAGGCAAGTGTCCATGACCACTAG 660  
DB 601 ACTTTGAGACATGAGAGTTTGAAGATGGAAGGCAAGTGTCCATGACCACTAG 660

QY 661 GGCATGCTGGCCTTCTTGTGATGTGAGTGTACTCGGCCACACTCAACACTTCCG 720  
DB 661 GGCATGCTGGCCTTCTTGTGATGTGAGTGTACTCGGCCACACTCAACACTTCCG 720  
QY 721 GGCAGGAACCCATTATCTCGGTAACTAGGGGCCCCACACTCATGAAAGACAGTAC 780  
DB 721 GGCAGGAACCCATTATCTCGGTAACTAGGGGCCCCACACTCATGAAAGACAGTAC 780  
QY 781 CTGGCCTTTGGCTCAAGCAACCTCTGTAAGGCTGCGGTAGTACTGAGACTGG 840  
DB 781 CTGGCCTTTGGCTCAAGCAACCTCTGTAAGGCTGCGGTAGTACTGAGACTGG 840  
QY 841 GGCAGCTTACAGGGGAGACAGACAAAGTCTACTTCTTCTTCAAGGAGCGGCACTGGAG 900  
DB 841 GGCAGCTTACAGGGGAGACAGACAAAGTCTACTTCTTCTTCAAGGAGCGGCACTGGAG 900  
QY 901 TCCGACTGTATTCGCGAGACAGTGTGGCTCGTGTGGCCCCGTGTCTGCAAGGCGCATATG 960  
DB 901 TCCGACTGTATTCGCGAGACAGTGTGGCTCGTGTGGCCCCGTGTCTGCAAGGCGCATATG 960  
QY 961 GGGGGCGACAGGACCTGTGAGAGGAAGTGAACAGCTTCTGTAAGGCGGCGGCTGGCATGC 1020  
DB 961 GGGGGCGACAGGACCTGTGAGAGGAAGTGAACAGCTTCTGTAAGGCGGCGGCTGGCATGC 1020  
QY 1021 TCTGCCCGCACTGGCAGCTTACTTCAACAGCTGACAGGCGATGCACACCTGACAGAC 1080  
DB 1021 TCTGCCCGCACTGGCAGCTTACTTCAACAGCTGACAGGCGATGCACACCTGACAGAC 1080  
QY 1081 ACCCTCTGGCACAACACACACTTCTTGGGGTTTTCAAGACAGTGGGGGTGACATGTAC 1140  
DB 1081 ACCCTCTGGCACAACACACACTTCTTGGGGTTTTCAAGACAGTGGGGGTGACATGTAC 1140  
QY 1141 CTGTGGGCGCATGTGTAGTACCAAGTGTGAAGAGATCAGCGGGGTGTTAGAGGCCCTAT 1200  
DB 1141 CTGTGGGCGCATGTGTAGTACCAAGTGTGAAGAGATCAGCGGGGTGTTAGAGGCCCTAT 1200  
QY 1201 AAGGATACCATGAGGAAGCCCAAGAAAGTGGACCGCTACACTGACCTTACCCAGCCT 1260  
DB 1201 AAGGATACCATGAGGAAGCCCAAGAAAGTGGACCGCTACACTGACCTTACCCAGCCT 1260  
QY 1261 CGGCGTGGCTGCTGCAATTAACAACCTGGCATCGGCGCAGGGCTACACACGCTCCCTGGAG 1320  
DB 1261 CGGCGTGGCTGCTGCAATTAACAACCTGGCATCGGCGCAGGGCTACACACGCTCCCTGGAG 1320  
QY 1321 CTACCGCAGACATCTCTCAACTTCTGTAAGAAACACCGCTGATGAGAGAGAGTGGGG 1380  
DB 1321 CTACCGCAGACATCTCTCAACTTCTGTAAGAAACACCGCTGATGAGAGAGAGTGGGG 1380  
QY 1381 CCTGGGTGAGCGCGCCCTGCTCTGTGAAGAAAGGACCACTTACCCACTGTGTGGC 1440  
DB 1381 CCTGGGTGAGCGCGCCCTGCTCTGTGAAGAAAGGACCACTTACCCACTGTGTGGC 1440  
QY 1441 GACCGGGTTACAGAGACTTATGAGGACACTTATACGTGCTGTCACTTGGCAGAGAGAC 1500  
DB 1441 GACCGGGTTACAGAGACTTATGAGGACACTTATACGTGCTGTCACTTGGCAGAGAGAC 1500  
QY 1501 GCGTGGCTGCTCAAGCTGTGAGCCTGGGGCCCTGGGTTCACTGATTAGAGAGCTGCAG 1560  
DB 1501 GCGTGGCTGCTCAAGCTGTGAGCCTGGGGCCCTGGGTTCACTGATTAGAGAGCTGCAG 1560  
QY 1561 CTGTTTGAACAGAGAGCCATGAAAGCCTGGTGTATCTCAGAGCAAGAGTGTCTTT 1620  
DB 1561 CTGTTTGAACAGAGAGCCATGAAAGCCTGGTGTATCTCAGAGCAAGAGTGTCTTT 1620  
QY 1621 GCGGCTCCCGGCTCAAGCTGTGAGCTGCGCGCGGCGGAGTGCATTAAGATGTGCTCC 1680  
DB 1621 GCGGCTCCCGGCTCAAGCTGTGAGCTGCGCGCGGCGGAGTGCATTAAGATGTGCTCC 1680  
QY 1681 TGTGACAGATGTGTCTGCGCGGAGACCCCTATTGGGCTGGAGGCTCAACACAGCCGC 1740  
DB 1681 TGTGACAGATGTGTCTGCGCGGAGACCCCTATTGGGCTGGAGGCTCAACACAGCCGC 1740





```

OY 421 GCGATCTCTGGAGAGCCCGGTGGAGAGAGACTGATGTATCCAGAAAAGGAGAAC 480
    |||||
DB 421 GCGATCTCTGGAGAGCCCGGTGGAGAGAGACTGATGTATCCAGAAAAGGAGAAC 480
OY 481 AACCAAGCCAGAGTCTCAACTTCATCCGCTCTCTGACGCCCTTCAATTCCTCCACTG 540
    |||||
DB 481 AACCAAGCCAGAGTCTCAACTTCATCCGCTCTCTGACGCCCTTCAATTCCTCCACTG 540
OY 541 TAGCTGTGGCACTACGCTTCCAGCCCAAGTGCACCTACGTCACATGCTTACCTTC 600
    |||||
DB 541 TAGCTGTGGCACTACGCTTCCAGCCCAAGTGCACCTACGTCACATGCTTACCTTC 600
OY 601 ACTTTGGAGCATGAGAGATTGAAGATGGAGGGCAAGTGTCCCTATGACCCAGCTAAG 660
    |||||
DB 601 ACTTTGGAGCATGAGAGATTGAAGATGGAGGGCAAGTGTCCCTATGACCCAGCTAAG 660
OY 661 GGGCATGCTGGCTCTTCTTGGATGGATGAGCTGTACTGCGCCACATCACAACCTTCTG 720
    |||||
DB 661 GGGCATGCTGGCTCTTCTTGGATGGATGAGCTGTACTGCGCCACATCACAACCTTCTG 720
OY 721 GGCACGGAACCCATTATCTGCGTACATGAGGGGCCCCACACCTCATGAGAGAGATAC 780
    |||||
DB 721 GGCACGGAACCCATTATCTGCGTACATGAGGGGCCCCACACCTCATGAGAGAGATAC 780
OY 781 CTGGCTTTTGGCTCAACGAACCTCACTTGTAGGCTCTGCTATGTACCTGAGAGTGTG 840
    |||||
DB 781 CTGGCTTTTGGCTCAACGAACCTCACTTGTAGGCTCTGCTATGTACCTGAGAGTGTG 840
OY 841 GGCAGCTTCACGGGGGAGACAGACAGACAGTCTACTTCTTTCAGAGGAGGGGAGTGGAG 900
    |||||
DB 841 GGCAGCTTCACGGGGGAGACAGACAGACAGTCTACTTCTTTCAGAGGAGGGGAGTGGAG 900
OY 901 TCCGACTGTATGCGCGAGAGTGTGGCTGTGTGGCCCGTGTCTGCAAGGGGAGATATG 960
    |||||
DB 901 TCCGACTGTATGCGCGAGAGTGTGGCTGTGTGGCCCGTGTCTGCAAGGGGAGATATG 960
OY 961 GGGGGGCGACGAGCCCTGACAGAGAGTGTGACACAGTCTTCTGAAAGGGCGGCTGGATGC 1020
    |||||
DB 961 GGGGGGCGACGAGCCCTGACAGAGAGTGTGACACAGTCTTCTGAAAGGGCGGCTGGATGC 1020
OY 1021 TCTGCCCCGAACTGGAGCTCTACTTCAACGCTGAGGGGAGTGCACACCTCTCAGAGAC 1080
    |||||
DB 1021 TCTGCCCCGAACTGGAGCTCTACTTCAACGCTGAGGGGAGTGCACACCTCTCAGAGAC 1080
OY 1081 ACCGCTGGGACAAACACACCTTCTTGGGTTTTCAGACAGTGGGGTGCATGTAC 1140
    |||||
DB 1081 ACCGCTGGGACAAACACACCTTCTTGGGTTTTCAGACAGTGGGGTGCATGTAC 1140
OY 1141 CTGTGGGCATCTGTGATGACAGTGTGAAGAGATCCAGCGGCTGTTTGAAGGGCCCTAT 1200
    |||||
DB 1141 CTGTGGGCATCTGTGATGACAGTGTGAAGAGATCCAGCGGCTGTTTGAAGGGCCCTAT 1200
OY 1201 AAGAGATACATGAGAGAGAGAGAGAGTGGAGCCGCTACACTGACCTGTACCCAGCCCT 1260
    |||||
DB 1201 AAGAGATACATGAGAGAGAGAGAGAGTGGAGCCGCTACACTGACCTGTACCCAGCCCT 1260
OY 1261 CGGCTGTGCTGTGATTAACAACCTGAGTGGAGGCGGACAGCTCCTCTGGAG 1320
    |||||
DB 1261 CGGCTGTGCTGTGATTAACAACCTGAGTGGAGGCGGACAGCTCCTCTGGAG 1320
OY 1321 CTACCCGACAAACATCTCAACTTGTCAAGAGACCCGCTGATGAGAGAGAGAGTGGGG 1380
    |||||
DB 1321 CTACCCGACAAACATCTCAACTTGTCAAGAGACCCGCTGATGAGAGAGAGAGTGGGG 1380
OY 1381 CTTGTGGTGGAGCCGCCCTGCTGTGTGAAGAGGACACAACTTCCACCTGCTGAGCC 1440
    |||||
DB 1381 CTTGTGGTGGAGCCGCCCTGCTGTGTGAAGAGGACACAACTTCCACCTGCTGAGCC 1440
OY 1441 GACCGGGTTACAGAGTGTGATGAGAGCCATATACAGTGTCTCATTTGGACAGAGAGAC 1500
    |||||
DB 1441 GACCGGGTTACAGAGTGTGATGAGAGCCATATACAGTGTCTCATTTGGACAGAGAGAC 1500

```

```

OY 1501 GCGTGGCTGCTCAAGGCTGTGAGCTTGGGGCCCTGGGTTTACCTGATTGAGAGCTGACG 1560
    |||||
DB 1501 GCGTGGCTGCTCAAGGCTGTGAGCTTGGGGCCCTGGGTTTACCTGATTGAGAGCTGACG 1560
OY 1561 CTGTTTGGACAGAGAGCCCATAGAAAGCTGTGTCTATCTTCAGAGACAGAGCTCTCTT 1620
    |||||
DB 1561 CTGTTTGGACAGAGAGCCCATAGAAAGCTGTGTCTATCTTCAGAGACAGAGCTCTCTT 1620
OY 1621 GCGGCTCCGCTGTCAGCTGTGAGCTGTGCGGCGGCGGAGCTGATTAATTCGCTCC 1680
    |||||
DB 1621 GCGGCTCCGCTGTCAGCTGTGAGCTGTGCGGCGGCGGAGCTGATTAATTCGCTCC 1680
OY 1681 TGTGACAGCTGTCTCTGCGCCCGGAGCCCTATTTGGCGCTGAGCGTCAACAGAGCCG 1740
    |||||
DB 1681 TGTGACAGCTGTCTCTGCGCCCGGAGCCCTATTTGGCGCTGAGCGTCAACAGAGCCG 1740
OY 1741 TGTGTGGCCGTGGGTGGGCACTTTGATCTTACTGATCCAGCATGTGATGACCTGGAG 1800
    |||||
DB 1741 TGTGTGGCCGTGGGTGGGCACTTTGATCTTACTGATCCAGCATGTGATGACCTGGAG 1800
OY 1801 ACTTCAGAGCATTTGCAACCTCCGTCGAGTAAATACAGTACAGGCGCCCTCCAAAAAC 1860
    |||||
DB 1801 ACTTCAGAGCATTTGCAACCTCCGTCGAGTAAATACAGTACAGGCGCCCTCCAAAAAC 1860
OY 1861 ATCACGGTGTGGGGGACAGACCTGTGCTGCGCTGCGCCACTCTCTCCACTTGGCC 1920
    |||||
DB 1861 ATCACGGTGTGGGGGACAGACCTGTGCTGCGCTGCGCCACTCTCTCCACTTGGCC 1920
OY 1921 CTG 1923
    |||
DB 1918 ATCACGGTGTGGGGGACAGACCTGTGCTGCGCTGCGCCACTCTCTCCACTTGGCC 1917
    |||
DB 1918 CGG 1920

```

```

RESULT 4
AX704763 2646 bp DNA linear PAT 04-APR-2003
LOCUS Sequence 67 from Patent W00202634.
DEFINITION AX704763
ACCESSION AX704763.1 GI:29561429
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulalia; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Gururajan, R., Hafalia, A.J., Kellick, D.A., Patterson, C., Azimzal, Y.,
Khan, F.A., Xu, Y., Iao, M.G., Yue, H., Au-Yang, U., Batra, S.,
Baughn, M.R., Borowsky, E.A., Lo, T.P., Lu, D.A., Lu, Y., Tang, Y.T.,
Yang, J., Zingler, K.A., Deleage, A.M., Gietzen, K., Marcus, G.A.,
Nguyen, D.B., Policky, J.L., Ramkumar, J., Thangavelu, K., Walle, N.K.
and Warren, B.A.
TITLE Human extracellular matrix and cell adhesion polypeptides
JOURNAL Patent: WO 0202634-A 67 10-JAN-2002;
INCYTE GENOMICS INC. (US)
FEATURES
source
1..2646
/oranism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/ncbi="Incyte ID No: 7950094CB1"
BASE COUNT 487 a 812 c 845 g 502 t
ORIGIN
Query Match 77.2% Score 1764.2 DB 6; Length 2646;
Best Local Similarity 97.5% Pred. No. 0;
Matches 1813; Conservatve 0; Mismatches 43; Indels 3; Gaps 2;
OY 84 CAGCGCGGAGACCGGACCTCGCTGCGAGCGCGGACAGACAGAGTGGGCTGAA 143
    |||||
DB 52 CCGAAGCGAGAGCGCGCGCGCGGAGCTGAGCTGCGGCAATGAGACAGTGGCTTAA 111
    |||||
OY 144 GCTGAGACCGGCGGCGTGGCGCATGGCCCAACTGGGCTGTGCTGCTGCGACAGAG 203
    |||||

```

Db 112 GCTCAGAGCCGGGCTGCGCATGAGCCCACTGGGCTGTCTGGCTGCTGGCAGCAAG 171  
Oy 204 GCTGAGGCGCTGGGCAATTTGGGCTGAGCTGTGTGGAACTTTGTGCCCTTAAGACACT 263  
Db 172 GCTGAGGCGCTGGGCAATTTGGGCTGAGCTGTGTGGAACTTTGTGCCCTTAAGACACT 231  
Oy 264 GCTTCTTGGGAGCTGGCCACGCTAGTACGGGGTTCTCCAGACGGGATCCAGACTT 323  
Db 222 GCTTCTTGGGAGCTGGCCACGCTAGTACGGGGTTCTCCAGACGGGATCCAGACTT 291  
Oy 324 CTTGACACTGACGCTGACGAGGAGCCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCTT 383  
Db 292 CTTGACACTGACGCTGACGAGGAGCCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCTT 351  
Oy 384 GTTTCCCTTACAGCATGAGAGGCGCTGAGCTGCAAGAGGAGATCTCTGGGAGCCCGCT 443  
Db 352 GTTTCCTTACAGCATGAGAGGCGCTGAGCTGCAAGAGGAGATCTCTGGGAGCCCGCT 411  
Oy 444 GGAGAAAGAAAGTGTATGTCAGAAAGGAGAAACCAACGAGCCGATGCTTCAACTT 503  
Db 412 GGAGAAAGAAAGTGTATGTCAGAAAGGAGAAACCAACGAGCCGATGCTTCAACTT 471  
Oy 504 CATCCGCTTCTGACAGCCCTACAAATGCCCTCCACCTGTACGTGTGGCACTACGCTT 563  
Db 472 CATCCGCTTCTGACAGCCCTACAAATGCCCTCCACCTGTACGTGTGGCACTACGCTT 531  
Oy 564 CCAGCCCAATGTCACCTTACATGTCACCTTCACTTGTGAGCATGAGAGTTTGA 623  
Db 532 CCAGCCCAATGTCACCTTACATGTCACCTTCACTTGTGAGCATGAGAGTTTGA 591  
Oy 624 AGATGGAAGGGAAGTGTCCCTATGACCCAGCTAAAGGSCATGGTGGCTTCTGTGA 683  
Db 592 AGATGGAAGGGAAGTGTCCCTATGACCCAGCTAAAGGSCATGGTGGCTTCTGTGA 651  
Oy 684 TGGTGAAGTGTACTGCGCCACACTCAACAATCTCTGGGACGAGAAACCATTAATCTGG 743  
Db 652 TGGTGAAGTGTACTGCGCCACACTCAACAATCTCTGGGACGAGAAACCATTAATCTGG 711  
Oy 744 TAACTTGGGGCCCAACCACTCCATGAGAGAGTACCTGGCTTTGGCTCAACGAAAC 803  
Db 712 TAACTTGGGGCCCAACCACTCCATGAGAGAGTACCTGGCTTTGGCTCAACGAAAC 771  
Oy 804 TCACTTTGAGGCTGTGGCTATGTACCTGAGAGTGTGGGACACTTACAGGGGAGAGAGA 863  
Db 772 TCACTTTGAGGCTGTGGCTATGTACCTGAGAGTGTGGGACACTTACAGGGGAGAGAGA 831  
Oy 864 CAAGGTCTACTTCTTCAAGGAGCGGCACTGTGACCTCCAGCTCTATGCCGACAGT 923  
Db 832 CAAGGTCTACTTCTTCAAGGAGCGGCACTGTGAGTCCGACTCTATGCCGACAGT 891  
Oy 924 GGTGGCTGTGTGGGCGCTGTCTGCAAGGGGCAATATGGGGGGCGACAGACCTTCAGAG 983  
Db 892 GGTGGCTGTGTGGGCGCTGTCTGCAAGGGGCAATATGGGGGGCGACAGACCTTCAGAG 951  
Oy 984 GAAAGTGAGCAAGTCTCTGAAGGCGGCTGGCATGTCTGGCCGCAACTGGCAGCTCTA 1043  
Db 952 GAAAGTGAGCAAGTCTCTGAAGGCGGCTGGCATGTCTGGCCGCAACTGGCAGCTCTA 1011  
Oy 1044 CTTCAACAGCTGACAGGCGATGACACCTTCGAGAGACACTTCTGGCAACACCACTT 1103  
Db 1012 CTTCAACAGCTGACAGGCGATGACACCTTCGAGAGACACTTCTGGCAACACCACTT 1071  
Oy 1104 CTTTGGGGTTTTTCAAGCAACACTGGGTGACATGTACTGTGGCCACTGTGACTACCA 1163  
Db 1072 CTTTGGGGTTTTTCAAGCAACACTGGGTGACATGTACTGTGGCCACTGTGACTACCA 1131  
Oy 1164 GTTGAAGAGATCCAGGGGTGTTGAGGGCCCTATAAGAGTACCATGAGAGAGGCCCA 1223  
Db 1132 GTTGAAGAGATCCAGGGGTGTTGAGGGCCCTATAAGAGTACCATGAGAGAGGCCCA 1191  
Oy 1224 GAAAGTGAGCAAGTCTCTGACCTGTACCCAGCCCTGGCTGTGCTGATTAACAA 1283  
Db 1192 GAAAGTGAGCAAGTCTCTGACCTGTACCCAGCCCTGGCTGTGCTGATTAACAA 1251

Oy 1284 CTGGCATGGGCGCACAGGGCTACACCAAGCTCCCTGAGCTACCCGACACATCTCAACTT 1343  
Db 1252 CTGGCATGGGCGCACAGGGCTACACCAAGCTCCCTGAGCTACCCGACACATCTCAACTT 1311  
Oy 1344 CGTCAAGAAAGCAACCGCTATGAGAGCAAGTGGGGCTCGGTGAGGCCGCCCTGCT 1403  
Db 1312 CGTCAAGAAAGCAACCGCTATGAGAGCAAGTGGGGCTCGGTGAGGCCGCCCTGCT 1371  
Oy 1404 CGTGAAGAGGACACCAACTTACACCACTGTGGGCGAGCCGGGTTTACAGACTTGATG 1463  
Db 1372 CGTGAAGAGGACACCAACTTACACCACTGTGGGCGAGCCGGGTTTACAGACTTGATG 1431  
Oy 1464 AGCCACTTATACAGTGTCTTATTTGACACAGAGAGAGGCTGGCTCAAGCTGTGAG 1523  
Db 1432 AGCCACTTATACAGTGTCTTATTTGACACAGAGAGAGGCTGGCTCAAGCTGTGAG 1491  
Oy 1524 CTTGGGGCCCTGGGTTCACTGATTTGAGAGAGTGCAGCTGTGTTGACAGAGACCTTGA 1583  
Db 1492 CTTGGGGCCCTGGGTTCACTGATTTGAGAGAGTGCAGCTGTGTTGACAGAGACCTTGA 1551  
Oy 1584 AAGCTGTGTATCTACAGAGCAAGAGTGTCTTGGCGGCTCCGCTCAGCTGTG 1643  
Db 1552 AAGCTGTGTATCTACAGAGCAAGAGTGTCTTGGCGGCTCCGCTCAGCTGTG 1611  
Oy 1644 GCAGCTGCCGCTGGCCGACTGCATTAAGATATGCTCTGTGACAGACTGTCTCGCCG 1703  
Db 1612 GCAGCTGCCGCTGGCCGACTGCATTAAGATATGCTCTGTGACAGACTGTCTCGCCG 1671  
Oy 1704 GGACCCCTATGCGCTGTGAGGCTGACACACAGCCGCTGTGTGGGCTGGCCACTT 1763  
Db 1672 GGACCCCTATGCGCTGTGAGGCTGACACACAGCCGCTGTGTGGGCTGGCCACTT 1731  
Oy 1764 TGGATCTTACTGATTCACAGCATGTGATGACTCGGACACTTACAGCAATTTGCAACTCG 1823  
Db 1732 TGGATCTTACTGATTCACAGCATGTGATGACTCGGACACTTACAGCAATTTGCAACTCG 1791  
Oy 1824 TGGCAATTAAGATACAGTCAAGGCCCACTNCCAAAACATACAGGTGGTGGGCGACAGA 1883  
Db 1792 TGGCAATTAAG--AAAGTCAGGCCCACT--CCCAAAAACATACAGGTGGTGGGCGACAGA 1848  
Oy 1884 CTTGGTGTGGCTGGCCGACTCTCTCCCAACTTGGGCGCCGCACTCCAAACCCGAGG 1942  
Db 1849 CTTGGTGTGGCTGGCCGACTCTCTCCCAACTTGGGCGCCGCACTTGGGG 1907

RESULT 5  
AX174731  
LOCUS AX174731 3293 bp DNA linear PAT 03-JUL-2001  
DEFINITION Sequence 25 from Patent WO0142285.  
ACCESSION AX174731  
VERSION AX174731.1 GI:14598254  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1. Yue, H., Tang, Y. T., Lal, P., Burford, N., Azimzal, Y., Patterson, C.,  
Baughn, M. R., Lu, D. A., Shah, P., and Au-Young, J.  
Extracellular matrix and cell adhesion proteins as well as genes  
encoding them  
Patent: WO 0142285-A 25 14-JUN-2001;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1..3293  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 1674661CB1"  
BASE COUNT 599 a 978 c 1052 g 664 t  
ORIGIN

Query Match	77.18	Score 1761	DB 6	Length 3293
Best Local Similarity	97.48	Pred. No. 0		
Matches 1811	Conservative 0	Mismatches 45	Indels 3	Gaps 2
QY	84	CAGGGCGGACCGGACCTCCCTCCACGCCGCCACACGACGACGACCTGGCCCTGAA	143	
DB	52	CCGAAACGGCAGGGCGGGCGGGGACCTGGAGCTGGCGCAATAGGACACCTGGCCCTGAA	111	
QY	144	GCTCAGAGCCGGGCGCTGCCCATGAGCCGCCACACACTGGGCTGTCTGGCTGTGGACAG	203	
DB	112	GCTCAGAGCCGGGCGCTGCCCATGAGCCGCCACACACTGGGCTGTCTGGCTGTGGACAG	171	
QY	204	GCTGTGGGCGCTGGGCTATTGGGCTGAGGTGTGTGGAACCTTGTGGCGGTGAAGACGT	263	
DB	172	GCTGTGGGCGCTGGGCTATTGGGCTGAGGTGTGTGGAACCTTGTGGCGGTGAAGACGT	231	
QY	264	GCTTCTGGGAGCTGGGCGCGGTGATAGCGCGGTCTCCAGACCGGCGATCCAGACCTT	323	
DB	232	GCTTCTGGGAGCTGGGCGCGGTGATAGCGCGGTCTCCAGACCGGCGATCCAGACCTT	291	
QY	324	CCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCGCT	383	
DB	292	CCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCGCT	351	
QY	384	GTTTGCCTTACGATGAGGCGCTGAGCTGCAAGAGCGATCTCTGGGAGGCGCCGCT	443	
DB	352	GTTTGCCTTACGATGAGGCGCTGAGCTGCAAGAGCGATCTCTGGGAGGCGCCGCT	411	
QY	444	GGAAGAAGACCTGATGTATCCAGAAAGGAAAGAACACAGACCGAGTGTTCACACTT	503	
DB	412	GGAAGAAGACCTGATGTATCCAGAAAGGAAAGAACACAGACCGAGTGTTCACACTT	471	
QY	504	CATCCGCTTCTCCAGCCCTACAAATGCTCCACCTGTACGTCTGTGGGACCTACGCGCT	563	
DB	472	CATCCGCTTCTCCAGCCCTACAAATGCTCCACCTGTACGTCTGTGGGACCTACGCGCT	531	
QY	564	CCAGCCCAAGTGCACCTACGTCAACATGCTCACCTTCACTTGGAGCATGAGAGTTGA	623	
DB	532	CCAGCCCAAGTGCACCTACGTCAACATGCTCACCTTCACTTGGAGCATGAGAGTTGA	591	
QY	624	AGATGGGAAGGGAAGTGTCCCTATGACCCAGTAAGGGGCAATGTGGCTTCTTGTGA	683	
DB	592	AGATGGGAAGGGAAGTGTCCCTATGACCCAGTAAGGGGCAATGTGGCTTCTTGTGA	651	
QY	684	TGTTGAGCTGTACTCGGCCACACTCAACACTTCTGTGGGACGGAACCACTTATCTGCG	743	
DB	652	TGTTGAGCTGTACTCGGCCACACTCAACACTTCTGTGGGACGGAACCACTTATCTGCG	711	
QY	744	TAAACATGGGGCCCAACACTCCATGAAGACAGATGCTGGCTTGTGCTCAACGAAC	803	
DB	712	TAAACATGGGGCCCAACACTCCATGAAGACAGATGCTGGCTTGTGCTCAACGAAC	771	
QY	804	TCACTTGTAGGCTGTGGCTATGTACCTGAGATGTGGGAGGCTGACGGGGGACGACGA	863	
DB	772	TCACTTGTAGGCTGTGGCTATGTACCTGAGATGTGGGAGGCTGACGGGGGACGACGA	831	
QY	864	CAAGTCTACTTCTTCTTACGAGGAGGCGAGTGCATGCTATCCGAGCAGGT	923	
DB	832	CAAGTCTACTTCTTCTTACGAGGAGGCGAGTGCATGCTATCCGAGCAGGT	891	
QY	924	GGTGGCTCGTGTGGCCCGTGTCTGCAAGGGCGATGTGGGGGCGACGACCTGTGACAG	983	
DB	892	GGTGGCTCGTGTGGCCCGTGTCTGCAAGGGCGATGTGGGGGCGACGACCTGTGACAG	951	
QY	984	GAAGTGACACGCTTCTGAAAGGCGGGGCGATGCTGTCCCGCAACCTGCAAGCTCTA	1043	
DB	952	GAAGTGACACGCTTCTGAAAGGCGGGGCGATGCTGTCCCGCAACCTGCAAGCTCTA	1011	
QY	1044	CTTCAACACGCTGACGAGCGATGCACACCTGACGACACTCTTGGGACAAACACACTT	1103	
DB	1012	CTTCAACACGCTGACGAGCGATGCACACCTGACGACACTCTTGGGACAAACACACTT	1071	
QY	1104	CTTGGGGTTTTCAAAGCAGTGGGGTGACATGTACTGTGGCAATCTGTGATGATACA	1163	

DB	1072	CTTGGGGTTTTCAAAGCAGTGGGGTGACATGTACTCTCGGCCATCTGTGATGATACA	1131	
QY	1164	GTTGGAAGATATCCAGGGGCTGTTTGAGGGCCCTTATAGAGTATACATAGAGAGCCCA	1223	
DB	1132	GTTGGAAGATATCCAGGGGCTGTTTGAGGGCCCTTATAGAGTATACATAGAGAGCCCA	1191	
QY	1224	GAAATGGAGCGGTACACTACCTGTATCCAGCCCTGGGCTGTGCTGTGATTAACA	1283	
DB	1192	GAAATGGAGCGGTACACTACCTGTATCCAGCCCTGGGCTGTGCTGTGATTAACA	1251	
QY	1284	CTGGCATGGGCGCACGGCTTACACACAGCTCCCTGTGACCTACCGCAACATCTCAACTT	1343	
DB	1252	CTGGCATGGGCGCACGGCTTACACACAGCTCCCTGTGACCTACCGCAACATCTCAACTT	1311	
QY	1344	CGTCAAGAACACCCCGCTGATGAGAGGACAGATGTGGGCGCTCGTGTGAGCCGCCCTGCT	1403	
DB	1312	CGTCAAGAACACCCCGCTGATGAGAGGACAGATGTGGGCGCTCGTGTGAGCCGCCCTGCT	1371	
QY	1404	CGTGAAGAAAGGACACCAACTTACACCCACTGTGTGGCGGACCGGGTTACAGGACTGTATGG	1463	
DB	1372	CGTGAAGAAAGGACACCAACTTACACCCACTGTGTGGCGGACCGGGTTACAGGACTGTATGG	1431	
QY	1464	AGCCACTATACAGTGTGTTCATTGGCACAGAGAGCGGTGGCTGCTCAAGGCTGTGAG	1523	
DB	1432	AGCCACTATACAGTGTGTTCATTGGCACAGAGAGCGGTGGCTGCTCAAGGCTGTGAG	1491	
QY	1524	CTGTGGGCGCTGGGTTACCTGATGTAGAGAGCTGACCTGTTTACAGAGACCCATGAG	1583	
DB	1492	CTGTGGGCGCTGGGTTACCTGATGTAGAGAGCTGACCTGTTTACAGAGACCCATGAG	1551	
QY	1584	AAGCTGTGTCTATCTCAGAGCAAGAGTGTCTTGTGGCGGCTCCGCTCTCAGCTGTGT	1643	
DB	1552	AAGCTGTGTCTATCTCAGAGCAAGAGTGTCTTGTGGCGGCTCCGCTCTCAGCTGTGT	1611	
QY	1644	GCAAGCTGCGCGTGCGGCACTGCAATMAATATGCTCTGTGACAGACTGTCTCTGCGCG	1703	
DB	1612	GCAAGCTGCGCGTGCGGCACTGCAATMAATATGCTCTGTGACAGACTGTCTCTGCGCG	1671	
QY	1704	GGAACCCCTATTGGCGCTGTGAGAGCGTCAACACGAGCGCTGTGTGGCCCTGGGTGGCACTT	1763	
DB	1672	GGAACCCCTATTGGCGCTGTGAGAGCGTCAACACGAGCGCTGTGTGGCCCTGGGTGGCACTT	1731	
QY	1764	TGATCTTCTACTGATCCAGATGTATGATGACCTGTGGACCTTACAGACTTGTGCAACTCCG	1823	
DB	1732	TGATCTTCTACTGATCCAGATGTATGATGACCTGTGGACCTTACAGACTTGTGCAACTCCG	1791	
QY	1824	TGGCAGTAAATACATGACGAGCCCACTNCCAAAACATCAGGGTGTGGCGGACAGA	1883	
DB	1792	TGGCAGTAAATACATGACGAGCCCACTNCCAAAACATCAGGGTGTGGCGGACAGA	1848	
QY	1884	CTGTGGCTGCGCGGCGACACTTCTCCCAACTTGGCCCTGCCGCACTCCCAACCCCGAGG	1942	
DB	1849	CTGTGGCTGCGCGGCGACACTTCTCCCAACTTGGCCCTGCCGCACTCCCAACCCCGAGG	1907	

RESULT 6  
 AX528275  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1  
 Spaderna,S.K., Tchernev,V., Liu,X., Shenoy,S., Spytek,R.,  
 Zernhusen,B., Patturajan,M., Taupier,R.J., Rastelli,L., Grosse,W.M.,  
 Szekeres,E.S., Alsbrook,J.T., Lepley,D.M., Shen,L., Burgess,C.E.,  
 Shinkets,R.A. and Padigaru,M.  
 Proteins and nucleic acids encoding same



JOURNAL Patent: WO 0206339-A 13 24-JAN-2002;

Curagen Corporation (US)

FEATURES

Location/Qualifiers

1. 3868

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 717 a 1106 c 1207 g 838 t

ORIGIN

Query Match 76.1%; Score 1738; DB 6; Length 3868;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1782; Conservative 0; Mismatches 35; Indels 3; Gaps 2;

OY 123 CAGCAGACAGCTGGCCCTGAAGCTCAGAGCCGGGCGTGCCTCATGTGCCCCACACTGGGC 182  
DB |||||||  
OY 410 CAGCAGACAGCTGGCCCTGAAGCTCAGAGCCGGGCGTGCCTCATGTGCCCCACACTGGGC 469  
DB |||||||  
OY 183 TGTCTGGCTGCTGGCAGCAAGGCTGTGGGCGCTGGGCAATTGGGCTGAGGTGTGTGCA 242  
DB TGTCTGGCTGCTGGCAGCAAGGCTGTGGGCGCTGGGCAATTGGGCTGAGGTGTGTGCA 529  
OY 243 CCTTGTGCGCGGTAAAGACAGTGTCTTCTGGGGAGCTGSCACGGTATGACGGGTTCTC 302  
DB |||||||  
OY 530 CCTTGTGCGCGGTAAAGACAGTGTCTTCTGGGGAGCTGSCACGGTATGACGGGTTCTC 589  
DB |||||||  
OY 303 CCAAGCCGGCATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTA 362  
DB CCAAGCCGGCATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTA 590  
OY 590 CCAAGCCGGCATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTA 649  
DB |||||||  
OY 363 CGTGGGCGCGGAGAGGCGCGTGTTCCTTCAGACATGAGAGCGCCGTGAGCTGCAGAGAGC 422  
DB CGTGGGCGCGGAGAGGCGCGTGTTCCTTCAGACATGAGAGCGCTGAGAGCTGCAGAGAGC 650  
OY 650 CGTGGGCGCGGAGAGGCGCGTGTTCCTTCAGACATGAGAGCGCTGAGAGCTGCAGAGAGC 709  
DB |||||||  
OY 423 GATCTCTGAGAGGCGCCCTGAGAGAGAGACTGATGTATCCAGAAAGGAGAAACAA 482  
DB GATCTCTGAGAGGCGCCCTGAGAGAGAGACTGATGTATCCAGAAAGGAGAAACAA 710  
OY 710 GATCTCTGAGAGGCGCCCTGAGAGAGAGACTGATGTATCCAGAAAGGAGAAACAA 769  
DB |||||||  
OY 483 CCAGACGAGAGTGTTCACACTTCCGCTTCTGACAGCCCTACAAATGCTCCACCTGTA 542  
DB CCAGACGAGAGTGTTCACACTTCCGCTTCTGACAGCCCTACAAATGCTCCACCTGTA 770  
OY 770 CCAGACGAGAGTGTTCACACTTCCGCTTCTGACAGCCCTACAAATGCTCCACCTGTA 829  
DB |||||||  
OY 543 CGTGTGTCGACACTAGCGCTTCCAGCCCAAGTGCACCTACGTCACAAATGCTCCACCTGTA 602  
DB |||||||  
OY 830 CGTGTGTCGACACTAGCGCTTCCAGCCCAAGTGCACCTACGTCACAAATGCTCCACCTGTA 889  
DB |||||||  
OY 603 TTTGGAGCATGAGAGTGTGAAGATGSGAAGGCGCAAGTGTCCCTATGACCCAGTAAAGG 662  
DB TTTGGAGCATGAGAGTGTGAAGATGSGAAGGCGCAAGTGTCCCTATGACCCAGTAAAGG 890  
OY 890 TTTGGAGCATGAGAGTGTGAAGATGSGAAGGCGCAAGTGTCCCTATGACCCAGTAAAGG 949  
DB |||||||  
OY 663 CCATGCTGCGCTTCTTGTGATGCTGAGCTGTACTCGGCCACACTCAACAACTTCTCTGG 722  
DB CCATGCTGCGCTTCTTGTGATGCTGAGCTGTACTCGGCCACACTCAACAACTTCTCTGG 950  
OY 950 CCATGCTGCGCTTCTTGTGATGCTGAGCTGTACTCGGCCACACTCAACAACTTCTCTGG 1009  
DB |||||||  
OY 723 CAGGGAACCCATTTATCCGCGTAACTATGSGGCGCCACACACTCCANTGAGAGAGTACT 782  
DB |||||||  
OY 1010 CAGGGAACCCATTTATCCGCGTAACTATGSGGCGCCACACACTCCANTGAGAGAGTACT 1069  
DB |||||||  
OY 783 GGCCTTTTGGCTCAACGAACTCACTTGTAGGCTCTGCTATGATGATGAGAGTGGG 842  
DB GGCCTTTTGGCTCAACGAACTCACTTGTAGGCTCTGCTATGATGATGAGAGTGGG 1070  
OY 1070 GGCCTTTTGGCTCAACGAACTCACTTGTAGGCTCTGCTATGATGATGAGAGTGGG 1129  
DB |||||||  
OY 843 CAGCTTCAAGGGGAGCAGACAAAGTCTACTTCTTTCAGAGAGCGGGAGTGTGATC 902  
DB CAGCTTCAAGGGGAGCAGACAAAGTCTACTTCTTTCAGAGAGCGGGAGTGTGATC 1130  
OY 1130 CAGCTTCAAGGGGAGCAGACAAAGTCTACTTCTTTCAGAGAGCGGGAGTGTGATC 1189  
DB |||||||  
OY 903 CGACTGCTATGCCAGCAGAGTGTGTGCTGTGTGGCCCGTGTCTGTCAAGGGCGATATGG 962  
DB |||||||  
OY 1190 CGCTGTGATATCCGAGAGGT 1249  
DB |||||||  
OY 963 GGGGCGACGAGACCTGACAGAGAAATGAGACAGCTTCTGTAAGGGCGGGCTGGGATGTCT 1022  
DB |||||||  
OY 1250 GGGGCGACGAGACCTGACAGAGAAATGAGACAGCTTCTGTAAGGGCGGGCTGGGATGTCT 1309  
DB |||||||

OY 1023 TGCCCCGAACCTGGCAGCTCTACTTTCACACAGCTGACGGCGATGACACACCTGACAGAC 1082  
DB |||||||  
OY 1310 TGCCCCGAACCTGGCAGCTCTACTTTCACACAGCTGACGGCGATGACACACCTGACAGAC 1369  
DB |||||||  
OY 1083 CTCTGGCAGCAACACACCTCTCTTTGGGTTTTTAAACACAGTGGGTGACATCTACT 1142  
DB CTCTGGCAGCAACACACCTCTCTTTGGGTTTTTAAACACAGTGGGTGACATCTACT 1370  
OY 1370 CTCTGGCAGCAACACACCTCTCTTTGGGTTTTTAAACACAGTGGGTGACATCTACT 1429  
DB |||||||  
OY 1143 GTCCGCACTGTGTAGTACACAGTTGGAAGATCCAGCGGGGTGTGTGAGGGCCCTATTA 1202  
DB GTCCGCACTGTGTAGTACACAGTTGGAAGATCCAGCGGGGTGTGTGAGGGCCCTATTA 1430  
OY 1430 GTCCGCACTGTGTAGTACACAGTTGGAAGATCCAGCGGGGTGTGTGAGGGCCCTATTA 1489  
DB |||||||  
OY 1203 GAGATACCATGAGAGAGCCAGAAAGTGGGACCGCTACACTGACCTGTATACCGACCTTCG 1262  
DB GAGATACCATGAGAGAGCCAGAAAGTGGGACCGCTACACTGACCTGTATACCGACCTTCG 1490  
OY 1490 GAGATACCATGAGAGAGCCAGAAAGTGGGACCGCTACACTGACCTGTATACCGACCTTCG 1549  
DB |||||||  
OY 1263 GCCTGGCTGTGTGATTAACAACTGTGCGGCGCCACGGCTTACACACCTCTCTGTGAGCT 1322  
DB GCCTGGCTGTGTGATTAACAACTGTGCGGCGCCACGGCTTACACACCTCTCTGTGAGCT 1550  
OY 1550 GCCTGGCTGTGTGATTAACAACTGTGCGGCGCCACGGCTTACACACCTCTCTGTGAGCT 1609  
DB |||||||  
OY 1323 ACCGACAAACATCTCAACTTCTGTCAGAGACACCGCTGATGAGAGAGAGTGGGCGC 1382  
DB ACCGACAAACATCTCAACTTCTGTCAGAGACACCGCTGATGAGAGAGAGTGGGCGC 1610  
OY 1610 ACCGACAAACATCTCAACTTCTGTCAGAGACACCGCTGATGAGAGAGAGTGGGCGC 1669  
DB |||||||  
OY 1383 TCGGTGAGCGCCCGCTGCTGTGAGAGAGAGGACCAACTTACACCACTGTGTGGCGA 1442  
DB TCGGTGAGCGCCCGCTGCTGTGAGAGAGAGGACCAACTTACACCACTGTGTGGCGA 1670  
OY 1670 TCGGTGAGCGCCCGCTGCTGTGAGAGAGAGGACCAACTTACACCACTGTGTGGCGA 1729  
DB |||||||  
OY 1443 CCGGCTTACAGAGCTTGAAGAGAGCCACTATACAGTGTGTTCAATGGCAGAGAGAGG 1502  
DB CCGGCTTACAGAGCTTGAAGAGAGCCACTATACAGTGTGTTCAATGGCAGAGAGAGG 1730  
OY 1730 CCGGCTTACAGAGCTTGAAGAGAGCCACTATACAGTGTGTTCAATGGCAGAGAGAGG 1789  
DB |||||||  
OY 1503 CTGCTGTCTCAAGGCTGTGAGGCTGSGGCGCTGAGTACCTGATTTGAGAGAGTCACT 1562  
DB CTGCTGTCTCAAGGCTGTGAGGCTGSGGCGCTGAGTACCTGATTTGAGAGAGTCACT 1790  
OY 1790 ATGGCTGTCAAGGCTGTGAGGCTGSGGCGCTGAGTACCTGATTTGAGAGAGTCACT 1849  
DB |||||||  
OY 1563 GTTTGACCAAGAGCCATGAGAGAGCTGTGTCTATCTCAGAGCAAGAAAGTGTCTTTGC 1622  
DB GTTTGACCAAGAGCCATGAGAGAGCTGTGTCTATCTCAGTGTGTCTCAGAGAGTGTCTTTGC 1850  
OY 1850 GTTTGACCAAGAGCCATGAGAGAGCTGTGTCTATCTCAGTGTGTCTCAGAGAGTGTCTTTGC 1909  
DB |||||||  
OY 1623 CGGCTCCCGCTCAGCTGAGTGTGACACTTCCGCTGCGCCACAGTCAAAAGTATCCGCTCTG 1682  
DB CGGCTCCCGCTCAGCTGAGTGTGACACTTCCGCTGCGCCACAGTCAAAAGTATCCGCTCTG 1910  
OY 1910 CGGCTCCCGCTCAGCTGAGTGTGACACTTCCGCTGCGCCACAGTCAAAAGTATCCGCTCTG 1969  
DB |||||||  
OY 1683 TGCAGACTGTGTCTCGCGCCGAGCCCTATTTGCGCTGAGAGCGTCAACACAGCGCGTG 1742  
DB TGCAGACTGTGTCTCGCGCCGAGCCCTATTTGCGCTGAGAGCGTCAACACAGCGCGTG 1970  
OY 1970 TGCAGACTGTGTCTCGCGCCGAGCCCTATTTGCGCTGAGAGCGTCAACACAGCGCGTG 2029  
DB |||||||  
OY 1743 TGTGGCCGTGGTGCCACTTTGATCTTTACTGATTCAGACATGTGATGACTCGGACAC 1802  
DB TGTGGCCGTGGTGCCACTTTGATCTTTACTGATTCAGACATGTGATGACTCGGACAC 2030  
OY 2030 TGTGGCCGTGGTGCCACTTTGATCTTTACTGATTCAGACATGTGATGACTCGGACAC 2089  
DB |||||||  
OY 1803 TTTAGGCAATTTGCAACCTCCGTCAGTAAAGTAAAGTCAAGGAGCCCACTCCCAAAAT 1862  
DB TTTAGGCAATTTGCAACCTCCGTCAGTAAAGTAAAGTCAAGGAGCCCACTCCCAAAAT 2090  
OY 2090 TTTAGGCAATTTGCAACCTCCGTCAGTAAAGTAAAGTCAAGGAGCCCACTCCCAAAAT 2146  
DB |||||||  
OY 1863 CACGCTGTGGGCGGAGCAGACCTGTGTGCTGCGCTGACACCTCTCTCAACTTGGCCT 1922  
DB CACGCTGTGGGCGGAGCAGACCTGTGTGCTGCGCTGACACCTCTCTCAACTTGGCCT 2147  
OY 2147 CACGCTGTGGGCGGAGCAGACCTGTGTGCTGCGCTGACACCTCTCTCAACTTGGCCT 2206  
DB |||||||  
OY 1923 GCCCGACTCCAAACCCGAGG 1942  
DB GCCCGACTCCAAACCCGAGG 2207  
OY 2207 TGCCCGCTGACCTTTGGGG 2226  
DB TGCCCGCTGACCTTTGGGG

RESULT 7  
AX528279 3112 bp DNA linear PAT 21-NOV-2002  
LOCUS AX528279  
DEFINITION Sequence 17 from Patent WO0206339.  
ACCESSION AX528279  
VERSION AX528279.1 GI:25172577  
KEYWORDS



SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Spaderna, S.K., Tchernev, V., Liu, X., Shenoy, S., Spytek, R.,  
Zerhusen, B., Palturalian, M., Taupier, R.J., Rastelli, L., Grosse, W.M.,  
Szekeles, E.S., Alsobrook, J.I., Lepley, D.M., Shen, L., Burgess, C.E.,  
Shimkets, R.A. and Padigar, M.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0206339-A 17 24-JAN-2002;  
Curagen Corporation (US)  
FEATURES  
source  
1. .3112  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 559 a 948 c 957 g 648 t  
ORIGIN  
Query Match 75.4%; Score 1722; DB 6; Length 3112;  
Best Local Similarity 97.4%; Pred. No. 0;  
Matches 1772; Conservative 0; Mismatches 45; Indels 3; Gaps 2;  
QY 123 CAGCAGACAGCTGGCCCTGAAGCTCAGAGCCGGGGCGTGGCCATGGCCCCACACTGGGC 182  
DB 61 CAGCAGACAGCTGGCCCTGAAGCTCAGAGCCGGGGCGTGGCCATGGCCCCACACTGGGC 120  
QY 183 TGTCTGGCTGCTGGCAGCAAGGCTGTGGGCGCTGGGCGATTGGGGCTGAGGTGTGTGGAA 242  
DB 121 TGTCTGGCTGCTGGCAGCAAGGCTGTGGGCGCTGGGCGATTGGGGCTGAGGTGTGTGGAA 180  
QY 243 CCTTGTGCGCGCTTAAGACAGTCTTCTGGGGAGCTGGCCAGCGTAGTACGGCGTCTC 302  
DB 181 CCTTGTGCGCGCTTAAGACAGTCTTCTGGGGAGCTGGCCAGCGTAGTACGGCGTCTC 240  
QY 303 CCAGACCGGCATCCAGAGCTTCTGACACTGACGCTGACGAGGAGCCCACTGGGCTTCTGA 362  
DB 241 CCAGACCGGCATCCAGAGCTTCTGACACTGACGCTGACGAGGAGCCCACTGGGCTTCTGA 300  
QY 363 CGTGGGGCGCCGAGAGGCCCTGTTGGCTTCAGCATGGAGGCCCTGGAGCTGCAGAGAGC 422  
DB 301 CGTGGGGCGCCGAGAGGCCCTGTTGGCTTCAGCATGGAGGCCCTGGAGCTGCAGAGAGC 360  
QY 423 GATCTCTGGAGAGCCGCCCTGGAGAGAGAGACTGATGATCCAGAAAGGAGAGACAA 482  
DB 361 GATCTCTGGAGAGCCGCCCTGGAGAGAGAGACTGATGATCCAGAAAGGAGAGACAA 420  
QY 483 CCAGACCGAGTCTTCAACTTCACTCGCTTCTGCAAGCCCTACAGTCCCTCCACTGTA 542  
DB 421 CCAGACCGAGTCTTCAACTTCACTCGCTTCTGCAAGCCCTACAGTCCCTCCACTGTA 480  
QY 543 CGTCTGGGACCTAGAGCCCTTCAGAGCCAGTGCACATAGTCAACATGTCTACACTTAC 602  
DB 481 CGTCTGGGACCTAGAGCCCTTCAGAGCCAGTGCACATAGTCAACATGTCTACACTTAC 540  
QY 603 TTTGGAGCATGAGAGTTTGAAGATGGGAGGCAAGTGTCCCTATGACCACAGTAAGG 662  
DB 541 TTTGGAGCATGAGAGTTTGAAGATGGGAGGCAAGTGTCCCTATGACCACAGTAAGG 600  
QY 663 CCATGCTGGCCCTTCTGGAGTGTGAGCTGTAAGTGGCCACACTCAACAACTTCTGGG 722  
DB 601 CCATGCTGGCCCTTCTGGAGTGTGAGCTGTAAGTGGCCACACTCAACAACTTCTGGG 660  
QY 723 CAGGAAACCATATATCTGCTGAACANTGGGGCCCACTCATGAAAGACAGATACCT 782  
DB 661 CAGGAAACCATATATCTGCTGAACANTGGGGCCCACTCATGAAAGACAGATACCT 720  
QY 783 GGCCTTTTGGCTCAAGAACTTCTAGAGCTGTGCTATGATAGTACAGAGTGTGG 842  
DB 721 GGCCTTTTGGCTCAAGAACTTCTAGAGCTGTGCTATGATAGTACAGAGTGTGG 780  
QY 843 CAGCTTCACGGGGGAGCAGACAAAGTCTACTTCTTCAGGAGAGCGGGCACTGAGTCTC 902

DB 781 CAGCTTCACGGGGGAGCAGACAAAGTCTACTTCTTCAGGAGAGCGGGCAGTGTGAGTCTC 840  
QY 903 CGACTGCTATGCCAGAGAGGTGGTGGTGGGCGCCGCTGCTGCAAGGCGCATATAGG 962  
DB 841 CGACTGCTATGCCAGAGAGGTGGTGGTGGGCGCCGCTGCTGCAAGGCGCATATAGG 900  
QY 963 GGGCGCAGGACCCCTGCAGAGAGAAAGTGGACACCTTCTTAAGAGCGGCGTGCATGCTC 1022  
DB 901 GGGCGCAGGACCCCTGCAGAGAGAAAGTGGACACCTTCTTAAGAGCGGCGTGCATGCTC 960  
QY 1023 TGGCCCGAAGTGGCAGCTTACTTCAACAGCTGCAAGCGATGCAACACCTGCAGAGAC 1082  
DB 961 TGGCCCGAAGTGGCAGCTTACTTCAACAGCTGCAAGCGATGCAACACCTGCAGAGAC 1020  
QY 1083 CTCTGCGCACAACACACCTTCTTGGGGGTTTTTCAAGCAGAGTGGGTGACATGTACT 1142  
DB 1021 CTCTGCGCACAACACACCTTCTTGGGGGTTTTTCAAGCAGAGTGGGTGACATGTACT 1080  
QY 1143 GTCGGCATCTGTGATGATCAGATTGGAAGATCAGCGGGTGTGTAAGGAGCCCTATTA 1202  
DB 1081 GTCGGCATCTGTGATGATCAGATTGGAAGATCAGCGGGTGTGTAAGGAGCCCTATTA 1140  
QY 1203 GGAATACCATGAGAGAACCCAGAAAGTGGACCGCTACACTGACCTGTACCAACCCCTCG 1262  
DB 1141 GGAATACCATGAGAGAACCCAGAAAGTGGACCGCTACACTGACCTGTACCAACCCCTCG 1200  
QY 1263 GCCGTGCTGCTGCTTAACAACATGTGGCATCGGGCGCACAGGCTACACAGCTCCCTGGAGCT 1332  
DB 1201 GCCGTGCTGCTGCTTAACAACATGTGGCATCGGGCGCACAGGCTACACAGCTCCCTGGAGCT 1260  
QY 1323 ACCCGACAAATCTCTCAACTTCTGCTCAAGAAGACACCCCTGATGGAAGAGAGAGTGGGGCC 1382  
DB 1261 ACCCGACAAATCTCTCAACTTCTGCTCAAGAAGACACCCCTGATGGAAGAGAGAGTGGGGCC 1320  
QY 1383 TCGGTGAGCGCGCCCTGCTGCTGTAAGAAAGGACCACTTCAACCACTGTGTGGCGGA 1442  
DB 1321 TCGGTGAGCGCGCCCTGCTGCTGTAAGAAAGGACCACTTCAACCACTGTGTGGCGGA 1380  
QY 1443 CCGGGTTACAGAACTTATGAGAGCCACTATACAGTGTCTGTCATTTGGCCAGAGAGAGCG 1502  
DB 1381 CCGGGTTACAGAACTTATGAGAGCCACTATACAGTGTCTGTCATTTGGCCAGAGAGAGCG 1440  
QY 1503 CTGGCTGCTCAAGGCTTGAAGCTGGGGAGCCCTGGGTGACCTGATGAGAGAGTGCAGCT 1562  
DB 1441 CTGGCTGCTCAAGGCTTGAAGCTGGGGAGCCCTGGGTGACCTGATGAGAGAGTGCAGCT 1500  
QY 1563 GTTTGACAGAGCCCATGAGAAAGCTGTGTCTATCTCAGAGCAAGAAAGTGTCTTTGC 1622  
DB 1501 GTTTGACAGAGCCCATGAGAAAGCTGTGTCTATCTCAGAGCAAGAAAGTGTCTTTGC 1560  
QY 1623 CGGCTCCCGCTCAGTGTGTGAGAGCGTGGCCCGCTGGCGAGTGCATTAAGATATGCTCTCTG 1682  
DB 1561 CGGCTCCCGCTCAGTGTGTGAGAGCGTGGCCCGCTGGCGAGTGCATTAAGATATGCTCTCTG 1620  
QY 1683 TGCAGACTGTGTCTCCCGCGGAGCCCTTATTTGGGCTGGAGGCTGTAACACAGCGCTG 1742  
DB 1621 TGCAGACTGTGTCTCCCGCGGAGCCCTTATTTGGGCTGGAGGCTGTAACACAGCGCTG 1680  
QY 1743 TGTGGCCGTGGGTGGCCACTTTTGAATCTTACTGATTCAGACANTGTGATGACTGTGGACAC 1802  
DB 1681 TGTGGCCGTGGGTGGCCACTTGTGATCTTACTGATTCAGACANTGTGATGACTGTGGACAC 1740  
QY 1803 TTTAGGCAATTTGCAACCTCGTGGCACTTAAGATATGATCAGAGCGCCACTTNCCTAAAAAAT 1862  
DB 1741 TTTAGGCAATTTGCAACCTCGTGGCACTTAAG--AAATCTAGGGCCACT--CCCAAAAAAAT 1797  
QY 1863 CAGGTTGTGGGGGAGCAGACCTGTGTGCTGCTGACCTGCTCTCAACTTGTGGCTCT 1922  
DB 1798 CAGGTTGTGGGGGAGCAGACCTGTGTGCTGCTGACCTGCTCTCAACTTGTGGCTCT 1857  
QY 1923 GCCCGACTCCAAACCCGAGG 1942  
11 11 11 11

DB 1858 TGCCCGCTGACACTTGGGG 1877

RESULT 8  
AX528277  
LOCUS AX528277  
DEFINITION Sequence 15 from Patent WO0206339.  
ACCESSION AX528277  
VERSION AX528277.1 GI:25172576  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Spaderne, S.K., Tchernev, V., Liu, X., Shenoy, S., Spytek, K.,  
Zehusen, B., Paturajan, M., Taupier, R.J., Raselli, L., Grose, W.M.,  
Szekeres, E.S., Alsobrook, J.I., Lepley, D.M., Shen, L., Burgess, C.E.,  
Shinkens, R.A. and Padigaru, M.  
Proteins and nucleic acids encoding same  
Patent: WO 0206339-A 15 JAN-2002;  
Curagen Corporation (US)

TITLE  
JOURNAL  
Location/Qualifiers  
1..2558  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

FEATURES  
SOURCE  
1..2558  
Location/Qualifiers

BASE COUNT 468 a 788 c 801 g 501 t

ORIGIN

Query Match 74.4%; Score 1699; DB 6; Length 2558;  
Best Local Similarity 97.3%; Pred. No. 0;  
Matches 1749; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

146 TCAGAGCCGGGGCGTGGCCATGAGCCCACTGGGCTGTGTGCTGGCAGCAAGGC 205  
DB 1 TCAGAGCCGGGGCGTGGCCATGAGCCCACTGGGCTGTGTGCTGGCAGCAAGGC 60

206 TGTGGGGCTTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 265  
DB 1 TGTGGGGCTTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 120

266 CTTCTGGGGAGCTGGCCAGGTAGTACGGCGGTTCCTCCAGCCGGCATCCAGAGCTCC 325  
DB 1 CTTCTGGGGAGCTGGCCAGGTAGTACGGCGGTTCCTCCAGCCGGCATCCAGAGCTCC 180

326 TGACACTGACCTGACGAGAGCCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCTTGT 385  
DB 1 TGACACTGACCTGACGAGAGCCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCTTGT 240

386 TTGCGCTTCAGCATGAGAGCCCTGAGCTGCAAGAGAGCGATCTCTGGAGGCGCCCGTGG 445  
DB 241 CTGCACTGAGGCGCTCTGTGTGGCTGCAAGGCTGTGATCTCTGGAGGCGCCCGTGG 300

446 AGAAGAACATGAGGTATCCAGAAAGGAGAAACAAACCAACCGAGTGTCAACTTCA 505  
DB 301 AGAAGAACATGAGGTATCCAGAAAGGAGAAACAAACCAACCGAGTGTCAACTTCA 360

506 TCCGCTTCCTGACAGCCCTACAAATGCTCCACCTGTACGTGTGGGACCTTACGCTTCC 565  
DB 361 TCCGCTTCCTGACAGCCCTACAAATGCTCCACCTGTACGTGTGGGACCTTACGCTTCC 420

566 AGCCCAACTGACCTACGTCAACATGTCTACCTTGTGAGAGTGAAG 625  
DB 421 AGCCCAACTGACCTACGTCAACATGTCTACCTTGTGAGAGTGAAG 480

626 ATGGGAAGGGAATGTCCCTATGACCAAGTAAGGGCATGCTGGCTTCTTGGATG 685  
DB 481 ATGGGAAGGGAATGTCCCTATGACCAAGTAAGGGCATGCTGGCTTCTTGGATG 540

686 GTGAGCTGTACTGGGACCACTCAACAACTTGTGGGAGGAGAACCACTTATCCGCGTA 745  
DB 541 GTGAGCTGTACTGGGACCACTCAACAACTTGTGGGAGGAGAACCACTTATCCGCGTA 600

746 ACATGGGGCCCCACCACTCCATGAAGACAGACTACCTGGCCCTTTGGCTCAACGAACCTC 805  
DB 601 ACATGGGGCCCCACCACTCCATGAAGACAGACTACCTGGCCCTTTGGCTCAACGAACCTC 660

806 ACTTGTAGGCTCTGCTATGACTGAGAGTGTGGGACGCTTCAAGGGGAGCAGCAGCA 865  
DB 661 ACTTGTAGGCTCTGCTATGACTGAGAGTGTGGGAGCTTCAAGGGGAGCAGCAGCA 720

866 AGCTCTACTTCTTCAAGGAGGAGGAGAGTGGAGTCCGACTGCTATGGCCAGCAGTGG 925  
DB 721 AGCTCTACTTCTTCAAGGAGGAGGAGAGTGGAGTCCGACTGCTATGGCCAGCAGTGG 780

926 TGGCTGTGTGGCCCGTGTCTGCAAGGGCGATATGGGGGCGCAGGACCTGCAGAGGA 985  
DB 781 TGGCTGTGTGGCCCGTGTCTGCAAGGGCGATATGGGGGCGCAGGACCTGCAGAGGA 840

986 AGTGACCACTCTCTGTAAGAGGGGCGGTGGCATGCTGTGCCGAACTGGAGCTTACT 1045  
DB 841 AGTGACCACTCTCTGTAAGAGGGGCGGTGGCATGCTGTGCCGAACTGGAGCTTACT 900

1046 TCAACCACTGACGAGATGACACCTGCAGACACCTGCGGCAACACCACTTCT 1105  
DB 901 TCAACCACTGACGAGATGACACCTGCAGACACCTGCGGCAACACCACTTCT 960

1106 TTGGGGTTTTCAGACACAGTGGGGTGAACATGTACCTGTGGCCATCTGTGATGACAGT 1165  
DB 961 TTGGGGTTTTCAGACACAGTGGGGTGAACATGTACCTGTGGCCATCTGTGATGACAGT 1020

1166 TGGAGATATCCAGGGGGTGTGGAGGCGCCCTATTAAGAGATACATAGGAAGCCCA 1225  
DB 1021 TGGAGATATCCAGGGGGTGTGGAGGCGCCCTATTAAGAGATACATAGGAAGCCCA 1080

1226 AGTGAGACCGCTACACTGACCTGTACCCAGCCCTGCGGCTGTGCTGATTAACAAT 1285  
DB 1081 AGTGAGACCGCTACACTGACCTGTACCCAGCCCTGCGGCTGTGCTGATTAACAAT 1140

1286 GGCATGCGGCGCAGGCTACACACAGCTCCCTGGAGCTACCCGACAAATCTTCAACTTGG 1345  
DB 1141 GGCATGCGGCGCAGGCTACACACAGCTCCCTGGAGCTACCCGACAAATCTTCAACTTGG 1200

1346 TCAAGAACACCGGATGAGAGGAGAGAGTGGGGCTCGGTGGAGCCGCCCCCTGCTCG 1405  
DB 1201 TCAAGAACACCGGATGAGAGGAGAGAGTGGGGCTCGGTGGAGCCGCCCCCTGCTCG 1260

1406 TGAAGAAAGGACCAACTTCAACCCCACTGTGTGGCGGACCGGGTTACAGACTTGTATGAG 1465  
DB 1261 TGAAGAAAGGACCAACTTCAACCCCACTGTGTGGCGGACCGGGTTACAGACTTGTATGAG 1320

1466 CCACCTATACAGTGTCTTCAATTTGGCACAGAGAGCGCTGCTGCTCAAGGCTGTGAGCC 1525  
DB 1321 CCACCTATACAGTGTCTTCAATTTGGCACAGAGAGCGCTGCTGCTCAAGGCTGTGAGCC 1380

1526 TGGGGCCCTGGGTACCTGATGTAGAGCTGACACTGTTTGAACAGAGCCCATGAGAA 1585  
DB 1381 TGGGGCCCTGGGTACCTGATGTAGAGCTGACACTGTTTGAACAGAGCCCATGAGAA 1440

1586 GCTGTGCTATCTCAGAGCAAGAGTGTCTTGGCGGCTCCGCTCTCAGCTGTGTC 1645  
DB 1441 GCTGTGCTATCTCAGAGCAAGAGTGTCTTGGCGGCTCCGCTCTCAGCTGTGTC 1500

1646 AGCTCCCGTGGCGAGCTGACATTAAGTATCGCTCTGTGACAGACTGTCTCGCCCGGG 1705  
DB 1501 AGCTCCCGTGGCGAGCTGACATTAAGTATCGCTCTGTGACAGACTGTCTCGCCCGGG 1560

1706 ACCCTATTTGGCGCTGAGAGCGCAACACACAGCCGCTGTGGCCCTGGGCTGGCCACTTGG 1765  
DB 1561 ACCCTATTTGGCGCTGAGAGCGCAACACACAGCCGCTGTGGCCCTGGGCTGGCCACTTGG 1620

1766 GATCTTACTGATCCAGCATGTATGACCTGTGAGCACTTCAAGGCACTTGAACACTCGCG 1825  
DB 1621 GATCTTACTGATCCAGCATGTATGACCTGTGAGCACTTCAAGGCACTTGAACACTCGCG 1680

OY 1826 GCACTAAGATACGTCAGGCCACTNCCCAAAAACATCAGCGTGTGGCGGACAGACC 1885  
|||||  
DB 1681 GAACTAAG--AAAGTCAGGCCACT--CCCAAAAACATCAGCGTGTGGCGGACAGACC 1737  
OY 1886 TGGTGTGGCCCTGCGACCTCTCCCAACTGTGGCCCTGCCGATCCAGCCCGAGG 1942  
|||||  
DB 1738 TGTGTGTGGCCCTGCGACCTCTCCCAACTGTGGCCCGATGCCGCTGTGGG 1794  
RESULT 9  
BD171174  
LOCUS BD171174 3776 bp DNA linear PAT 17-JAN-2003  
DEFINITION Novel gene and protein encoded thereby.  
ACCESSION BD171174  
VERSION BD171174.1 GI:27876986  
KEYWORDS WO 02052005-A/30.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3776)  
AUTHORS Ohara,O., Nagase,T. and Nakajima,D.  
TITLE Novel gene and protein encoded thereby  
JOURNAL Patent: WO 02052005-A 30 04-JUL-2002;  
KAZUSA DNA RESEARCH INSTITUTE,OSAMU OHARA,TAKAHIRO NAGASE, DAISUKE  
NAKAJIMA  
COMMENT OS Homo sapiens (human)  
PN WO 02052005-A/30  
PD 04-JUL-2002  
PF 20-DEC-2001 WO 2001JP011217  
PR 22-DEC-2000 JP 00P 389742  
PI OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA  
PC C12N15/12,C07K14/47  
CC Novel gene and protein encoded thereby  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
source 1..3776  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 696 a 1088 c 1191 g 801 t  
ORIGIN  
Query Match 70.88; Score 1617.6; DB 6; Length 3776;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1652; Conservative 0; Mismatches 19; Indels 3; Gaps 2;  
OY 269 CTGGGAGCTGTGGCCACGCTACGGGGGTTCTCCAGACCGGATCCAGACTTCTGA 328  
|||  
DB 494 CTGCAGAGCTGGCCACGCTAGTACGGGGTTCTCCAGACCGGATCCAGACTTCTGA 553  
OY 329 CACTGAGCTGTGGAGGAGCCCTGGGCTCTGTACGTGGGGGCGCGAGAGCCCTGTTG 388  
|||||  
DB 554 CACTGAGCTGTGGAGGAGCCCTGGGCTCTGTACGTGGGGGCGCGAGAGCCCTGTTG 613  
OY 389 CCTTCAGCATGAGAGCCCTGTGAGCTGCAGAGAGGATCTCCTGGAGGAGCCCGCTGAGA 448  
|||||  
DB 614 CCTTCAGCATGAGAGCCCTGTGAGCTGCAGAGAGGATCTCCTGGAGGAGCCCGCTGAGA 673  
OY 449 AGAAGACTGAGTGTATCCAGAAAGGAGAAACAACACGAGAGTGTCAACTTCATCC 508  
|||||  
DB 674 AGAAGACTGAGTGTATCCAGAAAGGAGAAACAACACGAGAGTGTCAACTTCATCC 733  
OY 509 GCTTCTGAGAGCCCTACAAATGCTCCACCTGTACGTGTGGGACCTAGAGCTTCCAGC 568  
|||||  
DB 734 GCTTCTGAGAGCCCTACAAATGCTCCACCTGTACGTGTGGGACCTAGAGCTTCCAGC 793  
OY 569 CCAAGTCCACTTACGTCAACATGTCTCACTTCTTGGAGCATGAGAGTTTGAAGATG 628  
|||||  
DB 794 CCAAGTCCACTTACGTCAACATGTCTCACTTCTTGGAGCATGAGAGTTTGAAGATG 853  
OY 629 GGAAGGCAAGTGTCTCTATGACCAAGCTAAGGGCCATGCTGGCCCTTGTGATGATG 688

|||||  
DB 854 GGAAGGCAAGTGTCTCTATGACCAAGCTAAGGGCCATGCTGGCCCTTGTGATGATG 913  
OY 689 AGCTTACTTGGGCGACATCAACAACACTTCCGTGGGAGAGGAACCATTTATCTGGGTACA 748  
|||||  
DB 914 AGCTTACTTGGGCGACATCAACAACACTTCCGTGGGAGAGGAACCATTTATCTGGGTACA 973  
OY 749 TGGGCGCCACCACTCCATGAAGAGAGTACCTGCTTTTGGCTCAACGAACCTCACT 808  
|||||  
DB 974 TGGGCGCCACCACTCCATGAAGAGAGTACCTGCTTTTGGCTCAACGAACCTCACT 1033  
OY 809 TTGTAGCTTGCCCTATGTATGACCTGTAGAGTGTGGGACCTTCACGGGGAGCAGCAAG 868  
|||||  
DB 1034 TTGTAGCTTGCCCTATGTATGACCTGTAGAGTGTGGGACCTTCACGGGGAGCAGCAAG 1093  
OY 869 TCTACTTCTTCTTCAAGGAGGCGGAGTGGAGTCCGATGCTATGCCAGACAGTGGTGG 928  
|||||  
DB 1094 TCTACTTCTTCTTCAAGGAGGCGGAGTGGAGTCCGATGCTATGCCAGACAGTGGTGG 1153  
OY 929 CTGCTGTGGCCGTGTGCAAGGCGATATGGGGGGCGACGGACCTGCAGAGGAAGT 988  
|||||  
DB 1154 CTGCTGTGGCCGTGTGCAAGGCGATATGGGGGGCGACGGACCTGCAGAGGAAGT 1213  
OY 989 GGACACGTTCTGTGAAGGCGGCGCTGGCATGCTGCCCGCAACTGGCAGCTTACTTCA 1048  
|||||  
DB 1214 GGACACGTTCTGTGAAGGCGGCGCTGGCATGCTGCCCGCAACTGGCAGCTTACTTCA 1273  
OY 1049 ACCAGCTCAGAGCGATGTGCACACCCCTGCAGAGACCTCTGTGCACAAACACACTTCTTG 1108  
|||||  
DB 1274 ACCAGCTCAGAGCGATGTGCACACCCCTGCAGAGACCTCTGTGCACAAACACACTTCTTG 1333.  
OY 1109 GGGTTTTTCAAGCAGCAGTGGGGTGACATGTACCTGTGGCCATCTGTAGTACCAAGTTGG 1168  
|||||  
DB 1334 GGGTTTTTCAAGCAGCAGTGGGGTGACATGTACCTGTGGCCATCTGTAGTACCAAGTTGG 1393  
OY 1169 AAGATATCCAGCGGCTGTTTGAAGGCGCCCTATAGAGTACCATGAGGAAGCCAGAAAT 1228  
|||||  
DB 1394 AAGATATCCAGCGGCTGTTTGAAGGCGCCCTATAGAGTACCATGAGGAAGCCAGAAAT 1453  
OY 1229 GGGACCGGTACATGACCTGATACCAAGCCCTGGGCGGCGGCTGGCTGGCTTAACTAAGTGC 1288  
|||||  
DB 1454 GGGACCGGTACATGACCTGATACCAAGCCCTGGGCGGCGGCTGGCTGGCTTAACTAAGTGC 1513  
OY 1289 ATCGGCGCAGGCGCTACACCAAGCTCCCTGAGAGTACCCGACAACTCTCAACTCTGTCA 1348  
|||||  
DB 1514 ATCGGCGCAGGCGCTACACCAAGCTCCCTGAGAGTACCCGACAACTCTCAACTCTGTCA 1573  
OY 1349 AGAAGCACCCGCTGATGAGAGAGAGAGTGGGGCTCGGTGAGAGCGCCGCTGCTGTGA 1408  
|||||  
DB 1574 AGAAGCACCCGCTGATGAGAGAGAGAGTGGGGCTCGGTGAGAGCGCCGCTGCTGTGA 1633  
OY 1409 AGAAGGAGCAACACTTCAACCAAGCTGTGGGGCGACCGGGTTACAGAGATTGATGAGAGCA 1468  
|||||  
DB 1634 AGAAGGAGCAACACTTCAACCAAGCTGTGGGGCGACCGGGTTACAGAGATTGATGAGAGCA 1693  
OY 1469 CCTATACAGTGTCTTCAATTGGCACAGAGAGAGCGCTGTGCTCAAGGCTGTGAGACTGG 1528  
|||||  
DB 1694 CCTATACAGTGTCTTCAATTGGCACAGAGAGAGCGCTGTGCTCAAGGCTGTGAGACTGG 1753  
OY 1529 GGCCCTGGGTTTCACTGATGATGAGAGCTGCAAGCTTTTGGACAGAGCCCATGAGAACCC 1588  
|||||  
DB 1754 GGCCCTGGGTTTCACTGATGATGAGAGCTGCAAGCTTTTGGACAGAGCCCATGAGAACCC 1813  
OY 1589 TGGTGTCTATCTCAAGGAGAAAGAGTGTCTTGGCGGCTGCCGCTCTCAAGTGTGTGACG 1648  
|||||  
DB 1814 TGGTGTCTATCTCAAGGAGAAAGAGTGTCTTGGCGGCTGCCGCTCTCAAGTGTGTGACG 1873  
OY 1649 TGCCCGTGGCGAGCTGATAAAGTATCGCTCTGTGCAGAGTGTCTCTCGCCCGGAGAC 1708  
|||||  
DB 1874 TGCCCGTGGCGAGCTGATAAAGTATCGCTCTGTGCAGAGTGTCTCTCGCCCGGAGAC 1933  
OY 1709 CCTATGCGCGCTGAGAGCGTCAACACACCGCGCTGTGGCGGCTGGGCGCACTTTGGAT 1768  
|||||



OY		1229	GGGAGCGCCTACAGTGCACCCCTGTACCAGACCCCTGGCGCTGCTGCATTAAACAATCTGGC	1288
Db		1454	GGGAGCGCCTACAGTGCACCCCTGTACCAGACCCCTGGCGCTGCTGCATTAAACAATCTGGC	1513
OY		1289	ATCGGCCCAACGGGTACACACAGCTTCCTGTGAGACTACCCGACMAATCCTCAACTTCGTCA	1348
Db		1514	ATCGGCCCAACGGGTACACACAGCTTCCTGTGAGACTACCCGACMAATCCTCAACTTCGTCA	1573
OY		1349	AGAAGCACCCGCTGATGGAGAGCAGTGAGGGCCCTGGTTGAGAGCCGCCCTCGTCTGTGA	1408
Db		1574	AGAAGCACCCGCTGATGGAGAGCAGTGAGGGCCCTGGTTGAGAGCCGCCCTCGTCTGTGA	1633
OY		1409	AGAAGGCCACCACTTACACCCACTGGTGGGCCGACACGGGTTACAGAGACTTATGAGAGCCA	1468
Db		1634	AGAAGGCCACCACTTACACCCACTGGTGGGCCGACACGGGTTACAGAGACTTATGAGAGCCA	1693
OY		1469	CCTATTACAGTCTCTTCAATTTGGACAGAGACGGCTGGCTGCTCAAGGCTGTGAGCCCTGG	1528
Db		1694	CCTATTACAGTCTCTTCAATTTGGACAGAGACGGCTGGCTGCTCAAGGCTGTGAGCCCTGG	1753
OY		1529	GGCCCTGGGTTACCTGATTGAGAGAGCTGCAGCTGTTTGACAGAGAGCCCATGAGAAACC	1588
Db		1754	GGCCCTGGGTTACCTGATTGAGAGAGCTGCAGCTGTTTGACAGAGAGCCCATGAGAAACC	1813
OY		1589	TGTGTGCTATCTCAAGACGAAAGCTGCTCTTTGGCCGGCTCCCCTCTACGTGTGTGACGC	1648
Db		1814	TGTGTGCTATCTCAAGACGAAAGCTGCTCTTTGGCCGGCTCCCCTCTACGTGTGTGACGC	1873
OY		1649	TGCCCGTGGCCGATGTGCATAAAGTAATGCGTCTGTGCAGACTGTGTCTCGCCGGGAGCC	1708
Db		1874	TGCCCGTGGCCGATGTGCATAAAGTAATGCGTCTGTGCAGACTGTGTCTCGCCGGGAGCC	1933
OY		1709	CCTATTGCGCGCTGAGAGCGTCAACACACGACCGCTGTGTGGCCGTGGGTGGCCACTTTGGAT	1768
Db		1934	CCTATTGCGCGCTGAGAGCGTCAACACACGACCGCTGTGTGGCCGTGGGTGGCCACTTTGGAT	1993
OY		1769	CTTATCTGATCTCACAGATGTGATGACCTGTGGACACTTACAGGCATTGTGCATCCGTGGCA	1828
Db		1994	CTTATCTGATCTCACAGATGTGATGACCTGTGGACACTTACAGGCATTGTGCATCCGTGGCA	2053
OY		1829	GTAGATGATACGTAGGGCCACTNCCCAAAAACATCAGGTGTGGCGGGCACACACCTGG	1888
Db		2054	GTAAG--AAATCAGGGCCACT-CCCAA AAAACATCAGGTGTGGCGGGCACACACCTGG	2110
OY		1889	TGCTGCCCTGGCACCTCTCTCCAACTTGGCCCTGTGCCGAGTCCAAACCCCGAG	1942
Db		2111	TGCTGCCCTGGCACCTCTCTCCAACTTGGCCGCAATGGCCGCAATGCCGCTTGAGACCTTTGGG	2164

RESULT 11  
AX535046                      8095 bp       DNA       linear    PAT 22-NOV-2002

AX535046                      Sequence 73 from Patent WO02068633.

LOCUS                          AX535046

DEFINITION                    AX535046.1 GI:25261698

ACCESSION                      Homo sapiens (human)

VERSION                        Homo sapiens

KEYWORDS                      Eukaryote; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

SOURCE                         Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
ORGANISM                        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL                         Meclina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.  
TITLE                            Compositions and methods relating to lung specific genes and  
proteins  
Patent: WO 02068633-A 73 06-SEP-2002:  
Diaderus, Inc. (US)  
FEATURES                        Location/Qualifiers  
source                            1..8095  
                                 /organism="Homo sapiens"  
                                 /mol\_type="genomic DNA"  
                                 /db\_xref="taxon:9606"

BASE COUNT                      1899 a       2246 c       2050 g       1900 t

ORIGIN

Query Match	Similarity	Score	DB	Length
Best Local Similarity	97.68%	Pred. No. 1,4e-277;		
Matches 1448;	Conservative	0;	Mismatches 30;	Indels 6; Gaps 3;
QY	462	TATCCACAAAAGGGAAGAACAACACGACCGAGTGTCTTCAACCTTCAATCCGCTTCTGACAGC	521	
Db	5887	TAAACGGACACACTTTCGTGCCCCCAGACCGAGTGTCTTCAACCTTCAATCCGCTTCTGACAGC	5946	
QY	522	CTACAAATGCTCCACCTGTACGTGTGTGGCACCTACGCTTCCAGCCCAAGTGCACCTA	581	
Db	5947	CTAAAGAGCCCTCCACCTGTACGTGTGTGGCACCTACGCTTCCAGCCCAAGTGCACCTA	6006	
QY	582	CGTAAACATGTGTCACCTTTCATTTGGAGCATATGAGAGATTTGAACATGGGAAGGGCAAGTG	641	
Db	6007	CGTAAACATGTGTCACCTTTCATTTGGAGCATATGAGAGATTTGAACATGGGAAGGGCAAGTG	6066	
QY	642	TCCCTATGACCCACACTAAGGGCCATATGCTGGCCCTTCTGTGGATGGTAGCTGTACTGCGC	701	
Db	6067	TCCCTATGACCCACACTAAGGGCCATATGCTGGCCCTTCTGTGGATGGTAGCTGTACTGCGC	6126	
QY	702	CACACTCAACACTTCTGCGGACAGGAACCCATTATCCTGCGTAACATGGGGCCCCACCA	761	
Db	6127	CACACTCAACACTTCTGCGGACAGGAACCCATTATCCTGCGTAACATGGGGCCCCACCA	6186	
QY	762	CTCCATGAAGAACAAGTACCTGGGCTTTGGCTCAAGAACCTCACTTTGTAGCTGTGC	821	
Db	6187	CTCCATGAAGAACAAGTACCTGGGCTTTGGCTCAAGAACCTCACTTTGTAGCTGTGC	6246	
QY	822	CTATGTACTGAGAGTGTGGGACACTTACCGGGGGAGCAGACACAAGGCTCTACTTCTTCT	881	
Db	6247	CTATGTACTGAGAGTGTGGGACACTTACCGGGGGAGCAGACACAAGGCTCTACTTCTTCT	6306	
QY	882	CAGGAGACGGGCACTGAGTCCACTGCTATGCCAGACAGTGTGTGCTGTGGCCCG	941	
Db	6307	CAGGAGACGGGCACTGAGTCCACTGCTATGCCAGACAGTGTGTGCTGTGGCCCG	6366	
QY	942	TGTTGGAAAGGGCATATAGGGGGGGGAGACGACCTGTAGAGGAAGTGCACACTTCT	1001	
Db	6367	TGTTGGAAAGGGCATATAGGGGGGGGAGACGACCTGTAGAGGAAGTGCACACTTCT	6426	
QY	1002	GAAAGCGCGGTGTCATATGCTGTGCCCGCAATGTGGCAGCTCTACTTCAACACAGCTGACAGC	1061	
Db	6427	GAAAGCGCGGTGTCATATGCTGTGCCCGCAATGTGGCAGCTCTACTTCAACACAGCTGACAGC	6486	
QY	1062	GATGCACACCTCTGACAGACACTCTCTGGCAGACACACACCTCTTGTGGGGTTTTCAGC	1121	
Db	6487	GATGCACACCTCTGACAGACACTCTCTGGCAGACACACACCTCTTGTGGGGTTTTCAGC	6546	
QY	1122	ACAGTGGGGGTGACATGTACCTGTGCGGCATGTGAGTACAGTGTGGAAAGATCCACGG	1181	
Db	6547	ACAGTGGGGGTGACATGTACCTGTGCGGCATGTGAGTACAGTGTGGAAAGATCCACGG	6606	
QY	1182	GGTGTGTTGAGGGCCCTATTAAGAGATACATAGAGAAAGCCAGAAAGTGTGGACCGCTACAC	1241	
Db	6607	GGTGTGTTGAGGGCCCTATTAAGAGATACATAGAGAAAGCCAGAAAGTGTGGACCGCTACAC	6666	
QY	1242	TGACCTGTACCAAGCCCTGTGGCTGTGCTGTGATTAACAATGCGATCGGCGCCACGG	1301	
Db	6667	TGACCTGTACCAAGCCCTGTGGCTGTGCTGTGATTAACAATGCGATCGGCGCCACGG	6726	
QY	1302	CTTACACACGCTCCCTGGAGCTATACCCGCAACATCTCAACTTCGTCAAGAAAGCACCCGT	1361	
Db	6727	CTTACACACGCTCCCTGGAGCTATACCCGCAACATCTCAACTTCGTCAAGAAAGCACCCGT	6786	
QY	1362	GATGAGAGAGAGTGGGGCTTCGCTGGAGAGCCGCGCCCTGTCTGTGAAGAAGGGCACCA	1421	
Db	6787	GATGAGAGAGAGTGGGGCTTCGCTGGAGAGCCGCGCCCTGTCTGTGAAGAAGGGCACCA	6846	
QY	1422	CTTACACCACTGTGTGGCCGACCGGGTTTACAGCACTTGATGAGACCACTTACAGTGT	1481	
Db	6847	CTTACACCACTGTGTGGCCGACCGGGTTTACAGCACTTGATGAGACCACTTACAGTGT	6906	

OY	1482	GTTCATTGGCAGAGAGAGCGGCTGGCTCTCAAGGCTGTAGGCGTGGGGCCCTGGGTTCA	1541
Db	6907	GTTCATTGGCAGAGAGAGCGGCTGGCTCTCAAGGCTGTAGGCGTGGGGCCCTGGGTTCA	6966
OY	1542	CCTGATTGAGAGGAGTGCAGCTGTTTGGACGAGAGCCCATGGAAGCCCTGATCTCA	1601
Db	6967	CCTGATTGAGAGGAGTGCAGCTGTTTGGACGAGAGCCCATGGAAGCCCTGATCTCA	7026
OY	1602	GAGCAAG---AAGCTGCTCTTGGCGGCTCCCGCTTCAGCTGGTGCAGCTGCCCCGTGGC	1658
Db	7027	GAGCAAGTTAAGCTGCTCTTGGCGGCTCCCGCTTCAGCTGGTGCAGCTGCCCCGTGGC	7086
OY	1659	CGACTGCATTAAGATGCTGCTCTGTGCGAGACGTGTCTCGCCCGGAGCCCTATTGGC	1718
Db	7087	CGACTGCATGAAGATGCTGCTCTGTGCGAGACGTGTCTCGCCCGGAGCCCTATTGGC	7146
OY	1719	CTGGAGGCTCAACACACACGCCCTGTGTGTGGCCGTGGTGGCCACTTTGATCTTACTGAT	1778
Db	7147	CTGGAGGCTCAACACACACGCCCTGTGTGTGGCCGTGGTGGCCACTTTGATCTTACTGAT	7206
OY	1779	CCAGCATGTGATGACCTCGGACACTTTCAGGCACTTTCGAACCTCCGTGGCACTAAGATCA	1838
Db	7207	CCAGCATGTGATGACCTCGGACACTTTCAGGCACTTTCGAACCTCCGTGGCACTAAG--AAA	7264
OY	1839	GTCAGGCCCACTTCCCAAAAACATCACGGTGGTGGCGGCGACAGACCTGTGTGCTGGCTG	1898
Db	7265	GTCAGGCCCACT--CCCAAAAACATCACGGTGGTGGCGGCGACAGACCTGTGTGCTGGCTG	7323
OY	1899	CCACTCTCTCTCCAACTTGGGCCCTGGCCGACATCCAAACCCGAGG	1942
Db	7324	CCACTCTCTCTCCAACTTGGGCCCAATGGCCCGCTGGACCTTTGGG	7367

[illegible]

OY	642	TCCCATATACCAGATAGAGGCATGTCGGCCCTTCTTGTGATGGTGGAGCTGTAATCGC	701
Db	6116	TCCCATATACCAGATAGAGGCATGTCGGCCCTTCTTGTGATGGTGGAGCTGTAATCGC	6175
OY	702	CACACTCAACAACCTTCTGTGGCCACGGAAACCATTATCTGGCTAACAATGGGGCCCAACA	761
Db	6176	CACACTCAACAACCTTCTGTGGCCACGGAAACCATTATCTGGCTAACAATGGGGCCCAACA	6235
OY	762	CTCCATGGAAGACAGATACCTTGCCCTTTGGCTCAAGCAACCTCACTTTGTAAGCTCGC	821
Db	6236	CTCCATGGAAGACAGATACCTTGCCCTTTGGCTCAAGCAACCTCACTTTGTAAGCTCGC	6295
OY	822	CTATGTACCTGAGAGTGTGGCAGCTTCAAGGGGACGACACAAGTCTACTTCTTCTT	881
Db	6296	CTATGTACCTGAGAGTGTGGCAGCTTCAAGGGGACGACACAAGTCTACTTCTTCTT	6355
OY	882	CAGGAGAGGGGACATGAGAGTCCGACATGTAATGCCAGAGAGTGGTGGCTGTGTGGCCG	941
Db	6356	CAGGAGAGGGGACATGAGAGTCCGACATGTAATGCCAGAGAGTGGTGGCTGTGTGGCCG	6415
OY	942	TGTCGCAAGGGGCGATATGGGGGGGCGACGACACCTCGACAGAGAAAGTGGACAGTTCCT	1001
Db	6416	TGTCGCAAGGGGCGATATGGGGGGGCGACGACACCTCGACAGAGAAAGTGGACAGTTCCT	6475
OY	1002	GAAGGCGCGGCTGGCATCTCTGTCCCGAATGGCAGCTTACTTCAACACAGCTGCAGGC	1061
Db	6476	GAAGGCGCGGCTGGCATCTCTGTCCCGAATGGCAGCTTACTTCAACACAGCTGCAGGC	6535
OY	1062	GATGCACACCCCTGCAGGACACCTCTCGCACACACACACCTTCTTTGGGGTTTTTCAMG	1121
Db	6536	GATGCACACCCCTGCAGGACACCTCTCGCACACACACACCTTCTTTGGGGTTTTTCAMG	6595
OY	1122	ACAAGTGGGGTACATGTACTGTGTGGCCATCTGTGAGTACCAAGTTGGAAAGATCCACCG	1181
Db	6596	ACAAGTGGGGTACATGTACTGTGTGGCCATCTGTGAGTACCAAGTTGGAAAGATCCACCG	6655
OY	1182	GCTGTTTAGGGGCCCCCTAATAGAGATACATAGAGAAAGCCAGAAAGTGGACCCGTACAC	1241
Db	6656	GCTGTTTAGGGGCCCCCTAATAGAGATACATAGAGAAAGCCAGAAAGTGGACCCGTACAC	6715
OY	1242	TGACCCTGTACCCACAGCCCTCGGCGCTGTGTGATTTAAACAATGGACATCGGCGCACGG	1301
Db	6716	TGACCCTGTACCCACAGCCCTCGGCGCTGTGTGATTTAAACAATGGACATCGGCGCACGG	6775
OY	1302	CTACACACAGCTCCCTGGAGCTACCCGACACATCCTCAACTTCTGTAAGAAAGACCCGCT	1361
Db	6776	CTACACACAGCTCCCTGGAGCTACCCGACACATCCTCAACTTCTGTAAGAAAGACCCGCT	6835
OY	1362	GATGAGAGAGAGAGTGGGGGCTCGGTTGGAGCGCGCCCTGCTCGTAAGAAAGGACACCA	1421
Db	6836	GATGAGAGAGAGAGTGGGGGCTCGGTTGGAGCGCGCCCTGCTCGTAAGAAAGGACACCA	6895
OY	1422	CTTCACCCACCTGGTGGCCGACCCGGGTTACAGACTTGATGAGACCACTATACAGTGTCT	1481
Db	6896	CTTCACCCACCTGGTGGCCGACCCGGGTTACAGACTTGATGAGACCACTATACAGTGTCT	6955
OY	1482	GTTCAATTGGCAGAGAGACGGCTGGTGGCTCAAGGCTGTGAGCCTGGGGGCCCTGGTTCA	1541
Db	6956	GTTCAATTGGCAGAGAGACGGCTGGTGGCTCAAGGCTGTGAGCCTGGGGGCCCTGGTTCA	7015
OY	1542	CCTGATTTGAGAGAGTGCAGTGTGTTTGACACAGAGCCCATGAAAGCCTGGTGCATCTCA	1601
Db	7016	CCTGATTTGAGAGAGTGCAGTGTGTTTGACACAGAGCCCATGAAAGCCTGGTGCATCTCA	7075
OY	1602	GAGCAAG---AAGTGTCTTTTGCCCGGCTCCGCTCTCAAGTGGTGCAGCTGCCCCGTGGC	1658
Db	7076	GAGCAAGTTAAGCTGTCTTTTGCCCGGCTCCGCTCTCAAGTGGTGCAGCTGCCCCGTGGC	7135
OY	1659	CGACTGCATAAAGTATGCGCTCTGTGCAGATGTGTCTCTCCCGGGAGACCCCTATGGCG	1718
Db	7136	CGACTGCATAAAGTATGCGCTCTGTGCAGATGTGTCTCTCCCGGGAGACCCCTATGGCG	7195

QY 1719 CTGAGCGTCACACACGCGCGTGTGTGGCCGTGGCCACTTGGATCTTACTGAT 1778  
 Db 7196 CTGAGCGTCACACACGCGCGTGTGTGGCCGTGGCCACTTGGATCTTACTGAT 7255  
 QY 1779 CCAGCATGTGATGACCTCGGACACTTTCAGGATTTGACACCTCGGTGAGTAAATACA 1838  
 Db 7256 CCAGCATGTGATGACCTCGGACACTTTCAGGATTTGACACCTCGGTGAGTAAAG--AAA 7313  
 QY 1839 GTCAGGCCCATTCGCCAAAACATCAGGTGTGTGGCGGACAGACCTGTGTGCTGCTG 1898  
 Db 7314 GTCAGGCCCATTCGCCAAAACATCAGGTGTGTGGCGGACAGACCTGTGTGCTGCTG 7372  
 QY 1899 CCACCTCTCTCCCACTTGGCCCTGCGCCGACTCCACACCCGAGG 1942  
 Db 7373 CCACCTCTCTCCCACTTGGCCCTGCGCCGACTCCACACCCGAGG 7416  
 RESULT 13  
 S79463 3503 bp mRNA linear ROD 09-DEC-1995  
 LOCUS M-Sema F-a factor in neural network development [mice, neonatal  
 DEFINITION brain, mRNA, 3503 nt].  
 ACCESSION S79463  
 VERSION S79463.1 GI:1110598  
 KEYWORDS  
 SOURCE Mus sp.  
 ORGANISM Mus sp.  
 REFERENCE 1 (bases 1 to 3503)  
 AUTHORS Inagaki,S., Furuyama,T. and Iwahashi,Y.  
 TITLE Identification of a member of mouse semaphorin family  
 JOURNAL FEBS Lett. 370 (3), 269-272 (1995)  
 MEDLINE 95385809  
 PUBMED 7656991  
 REMARK GenBank staff at the National Library of Medicine created this  
 entry [NCBI g1bseq 170690] from the original journal article.  
 This sequence comes from Fig. 1A.  
 FEATURES  
 source  
 1..3503  
 /organism="Mus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10095"  
 1..3503  
 /gene="M-Sema F"  
 /note="a factor in neural network development"  
 100..2604  
 /gene="M-Sema F"  
 /note="semaphorin homolog; This sequence comes from Fig.  
 1A"  
 /codon\_start=1  
 /protein\_id="AA035184.1"  
 /db\_xref="GI:1110598"  
 /translation="MAPHRAVWLLAAGLKGIGICAEEMMNIVPKRYVSGELVTVVR  
 FSTGIDFLFVLTGHSGLIVGARALAFVSEALELDQALISWEPAERKICTOK  
 GKSQNTCEFNIRFLQPNNSHLVCGTFAVQPCYITMILFTLDAEFDGKCP  
 YDPAKGTGLVDDELXYSATLNLGLTEPLRLRMTGHSKTYLAFLMELPHEVS  
 AFDVESGFTGDDDKLYFFSERAVEDYSEOVVAVRACVGDGAGATTLKMT  
 TPLKARLVGAPDMKVFENOLKAVTLRGSMTTFEGYVQARMGMDLSAVCEYOL  
 EJOIOVREGPYKEXSEDAOKMARTDPVPSRPSCTINNHRDNGTSSLEPNTLN  
 FIKHPLMEQVFKRLGRPLVKNKNTFTHVADRVGLDADATVTLFVIGDWMLK  
 AVSLGPWIHAVEELQVPDQPEVESLVLSQSKVLPASRSQVLQSLADCKYRRCVD  
 CVLARDPYCANVNTSRVATSGRSGLFVOHVAANDTSMCKQYGIKKYRSIPKNI  
 TVVSGDLVPLCHLSNMLAHMTGSQDLEAEOGSEFLVDTGOLAVYMAAOSRHS  
 PYRCYSEOGTRLAESYLVAVVAGSSVTEARAPLENGLWMLAVVALAGVCLLL  
 LVLSIRRLREELKGAASERLLVPLLEPKPASPFRPGPTDRLMDPVYYS  
 DGSIKIVGHAKROPGGGPPSPPGITGQDLPSTRLHLGGSRNSNANGVRLDGGG  
 DRGSGPLDELDELRRKLQKROPDPDSNPESSV"  
 BASE COUNT 726 a 917 c 1057 g 803 t  
 ORIGIN  
 Query Match 59.2%; Score 1351; DB 10; Length 3503;  
 Best Local Similarity 84.7%; Pred. No. 1,1e-266;

Matches 1539; Conservative 0; Mismatches 271; Indels 6; Gaps 2;  
 QY 127 AGGACAGTGGCCCTGAAGCTCAGAGCCGGGGGTGCGCCATGGCCCAACATGGGCTGTG 186  
 Db 61 AGGCTGTGTGACATGAAGTTTAAACCCCTGGGTGTCTGTGCATGGGCCCACTATGGGCTGTG 120  
 QY 187 TGGCTGTGTGACATGAAGCTGTGGGGCTGGGCAATGGGGCTGAGGTGTGTGAACCTT 246  
 Db 121 TGGCTGTGTGACATGAAGCTGTGGGGCTGGGCAATGGGGCTGAGGTGTGTGAACCTT 180  
 QY 247 GTGGCGGTAAGACAGTCTTCTTGGGAGAGTGTGGCCACAGTAAATAGCGGGTCTCCAG 306  
 Db 181 GTGGCGGTAAGACAGTCTTCTTGGGAGAGTGTGGCCACAGTAAATAGCGGGTCTCCAG 240  
 QY 307 ACCGCGATCCAGAGCTTCTGACACTGACGTGACGAGAGCCAGTGGGCTCTGTATGTG 366  
 Db 241 ACAGGCAATCCAGAGCTTCTGACACTGACGTGACGAGAGCTGTGATATGTG 300  
 QY 367 GGGCGCCGAGAGGCCCTGTGTTGCTTACAGATGAGAGGCCCTGAGCTGCAGAGAGCCATC 426  
 Db 301 GGGCGCCGAGAGGCCCTGTGTTGCTTACAGATGAGAGGCCCTGAGCTGCAGAGAGCCATC 360  
 QY 427 TCCTGGGAGGCCCCCGTGGAGAAAGACTAGTGTATCCAGAAAGGAAAGAACACAG 486  
 Db 361 TCCTGGGAGGCCCCCGTGGAGAAAGACTAGTGTATCCAGAAAGGAAAGAACACAG 420  
 QY 487 ACCGAGTCTTCACATTCATCCGCTCTGAGAGCCCTCAATGGCTCCACCTGTATGTG 546  
 Db 421 ACCGAGTCTTCACATTCATCCGCTCTGAGAGCCCTCAATGGCTCCACCTGTATGTG 480  
 QY 547 TGTGGACACTTACGCTTCCAGCCCAAGTGACACTTACGTCAACATGCTTCACTTGTG 606  
 Db 481 TGTGGACACTTACGCTTCCAGCCCAAGTGACACTTACGTCAACATGCTTCACTTGTG 540  
 QY 607 GAGCATGAGAGTGTGAAGATGGGAGGAGGCAAGTCTCCATAGACCCAGCTAAGGCGCAT 666  
 Db 541 GAGCATGAGAGTGTGAAGATGGGAGGAGGCAAGTCTCCATAGACCCAGCTAAGGCGCAT 600  
 QY 667 GCTGGCTCTTGTGATGTGAGCTGTGACCTGCGCACACTCAACATCTTCTGGGACG 726  
 Db 601 ACCGAGCTCTTGTGAGAGGTGTGAGCTGTGACCTGCGCACACTCAACATCTTCTGGGACG 660  
 QY 727 GAACCCATTATCTGCGTAAACATGAGGGGCCCAACACTTCATGAAGACAGATTAAGCTTGGCC 786  
 Db 661 GAGCCGTTATCTTGTGATGATGAGGGGCCCAACACTTCATGAAGACAGATTAAGCTTGGCC 720  
 QY 787 TTTTGGCTCAACGAACTCTTGTGAGGCTGTGCTATGTACCTGTGAGAGTGTGGGACG 846  
 Db 721 TTTTGGCTCAACGAACTCTTGTGAGGCTGTGCTATGTACCTGTGAGAGTGTGGGACG 780  
 QY 847 TTCAGGGGGAGACGACGACAGAGTCTACTTCTTCAAGGAGCGGGGAGTGGAGTCCGAC 906  
 Db 781 TTCAGGGGGAGACGACGACAGAGTCTACTTCTTCAAGGAGCGGGGAGTGGAGTCCGAC 840  
 QY 907 TGTATGCGGAGCAGAGTGTGGCTGTGTCGTCGTCGAAAGGCGGATATGGGGGCG 966  
 Db 841 TGTATGCGGAGCAGAGTGTGGCTGTGTCGTCGTCGAAAGGCGGATATGGGGGCG 900  
 QY 967 GCAGGAGACCTGCGAGAGAAATGTGACGACGTTCTGAAAGGCGGGCTGTGATGCTTCC 1026  
 Db 901 GCAGGAGACCTGCGAGAGAAATGTGACGACGTTCTGAAAGGCGGGCTGTGATGCTTCC 960  
 QY 1027 CCGAAGTGGAGCTTACTTCAACGAGTGCAGGAGTGCACACCTCTCAGACACCTCC 1086  
 Db 961 CCGAAGTGGAGCTTACTTCAACGAGTGCAGGAGTGCACACCTCTCAGACACCTCC 1020  
 QY 1087 TGGCAACAACACCTCTTGTGGGCTTTTCAAGCAGAGTGGGGTGCATGTATCTGTG 1146  
 Db 1021 TGGCAACAACACCTCTTGTGGGCTTTTCAAGCAGAGTGGGGTGCATGTATCTGTG 1080  
 QY 1147 GCCATCTGTGATACCACTTGGAAAGATCCAGAGGGGTGTGAGGGCCCTATTAAGAG 1206  
 Db 1081 GCCATCTGTGATACCACTTGGAAAGATCCAGAGGGGTGTGAGGGCCCTATTAAGAG 1140



OY		1207	TACCACTGAGGAAGCCACCAGAATGGGAGACCGGTACAACACTGCACCTGTACACCGCCTCGGCT	1266
Dd		1141	TACAGTAGACAAAGCCACAGAAATGGGCCCGCATACTATGACCCGGTACACCGCCTTGCGCT	1200
OY		1267	GCGTCGTGCATTAAACAACTGGCCATTCGCGGCACGAGGCTAACACAGCTCCCTGGAGTAACC	1326
Dd		1201	GGTTCCGTGTATCAACAACTGGCCACCGACAGACAATGCTACACAGTCCCTGGAACTGGCG	1266
OY		1387	TGGAAGCCCCCCCCTGCTGCTGTGAAGAAGGGGACACAACTTCACCCACCTGTGGTGGCGACCG	1446
Dd		1321	TTGGGCGCGCCCTACTCTGTGTAAAGAAACACTAATCTCACACACAGTGTGTGGCGACAG	1380
OY		1447	GTTACAGAGACTTWTAGAGACCACCTATAACAGTGTTCATTGTGGCACAGAGAGCGCTGG	1506
Dd		1381	GTCCCAAGCGGCTWTAGTGTGCGCCACCTATACAGTGTTCATTGTGTACAGAGAGATGGCTGG	1440
OY		1507	CTGCTCAAGGCGTGTGAGCGCTGGGGCCCTGGGCTTCAACCTGATTGAGAGAGCTGACTGTT	1566
Dd		1441	CTGCTGAAGCGTGTGAGCGCTGGGGCCCTGGATTCACATGATGTGAGAGAACTCCAGCTGTTT	1500
OY		1567	GACCAAGAGCCCATATGACAGAACCGCTGTCTATCTCAGAGACAAGACGTCTTTGGCGGC	1626
Dd		1501	GACCAAGAGCCCATGTGAGAAATGTCTGTGTCTCTCAGAGCAAAGAGTCTTTTGTCTGGC	1560
OY		1637	TCCCGCTCTCAGCTGTGTGACAGCTCCCGCTGGCGCCACATGACATAAAGTATCGCTGTGCA	1686
Dd		1561	TCCCGCTCTCAGCTGTGTGACAGCTCTCTCTGCGCCACATGCAACAAAGTACCGTTTCTGTGA	1620
OY		1687	GACTGTGTGCTCCGCGCCGGGAGACCCCTAATGTGGCGCTGGAGGCGTCAACACCGCGCTGTGTG	1746
Dd		1631	GACTGTGTGCTCCGCGCCGGGAGACCCCTTACTGTGTCTCTGGAATGTCAACACCGCGCTGTGTG	1680
OY		1747	G---CCGTGGGTGGCCACTTTTGATCTTACTGATCCAGCATGTGATGACCTCGGACACT	1803
Dd		1681	GCCACACACAGATGTGTGCGCTCGGGGCTCTTGTGGCCAATGTGGGCAACTGTGACACT	1740
OY		1804	TCAGGCAATTGCCAACCCTCCGTGGCAGTAAAGTAAACAGTACAGGCCCCACATNCCCAAAAACATC	1863
Dd		1741	TCCAAAGATGTGTATACCAAGTATGAGTAA--AAAAGTCAATATTAATTCACMAAACATC	1797
OY		1864	ACGGTGTGTGGGGGACACAGACCTGATGTGTGCTGTGCACACCTCTCTCAACTTGGGCCCTG	1923
Dd		1798	ACCGTGTGTGTAGGACAGACCTGTCTTACCTGTGCCACCTCTGTGTCCAAATTTGGCCCAT	1857
OY		1924	CCCGACTTCCAAACCCCG	1939
Dd		1858	GCCCACTGTGACCTTCG	1873
RESULT 14				
LOCUS		BC017476		
DEFINITION		Homo sapiens sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4c, mRNA (CDNA clone MGC:15189 IMAGE:3528227), complete cds.		
ACCESSION		BC017476		
VERSION		BC017476.1		
KEYWORDS		MGC.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS		Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenman,C.M., Shale,N.K., Altschul,S.F., Zeeberg,B., Bietlow,K.H., Schmeider,C.F., Hsieh,F., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J.J., Hsieh,F., Datchenko,L., Marsina,K., Farmer,A.A., Rubin,G.W., Hong,L.,		

TITLE	JOURNAL	REFERENCE	AUTHORS	COMMENT
STREPLTON, M., SOARES, M.B., DONALDO, M.F., CASARANI, T.L., SCHNEIDER, T.E., BROWNSTEIN, M.J., USDI, T.B., TOSHIYUKI, S., CARINCI, P., PRANGE, C., RAHA, S.S., LOQUELLANO, N.A., PETERS, G.J., ABRAMSON, R.D., MULLANY, S.J., BOSCH, S.A., MCEWAN, P.J., MCHERRAN, R.D., MALEK, J.A., GUNARATNE, P.H., RICHARDS, S., WOLLEY, K.C., HALE, S., GARCIA, A.M., GAY, L.J., HULIK, S.W., VALLADON, D.K., MUZY, D.M., SODERGREN, E.J., LU, X., GIBBS, R.A., FAHEY, J., HELTON, E., KETTEMAN, M., YODAN, A.C., RODRIGUES, S., SANCHEZ, A., WHITING, M., MEDAN, A., MOG, A.C., SHARVCHENKO, Y., BOUFFARD, G.G., BLAKESLEY, R.W., TOUCHMAN, J.W., GREEN, E.D., DICKSON, M.C., RODRIGUEZ, A.C., GLIMWOOD, J., SCHMITZ, D., MYERS, R.M., BUTTERFIELD, Y.S., KRZYWINSKI, M.I., SKALSKA, U., SMALLIS, D.E., SCHMERCH, A., SCHEIN, J.E., JONES, S.J. and MARRA, M.A.	Genet. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)	2 (bases 1 to 2907)	Strausberg, R. Direct Submission Submitted (19-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	NIR-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada Info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Ffeli, Erin Garland, Ran Guin, Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandon, Anna-Liisa Pradhu, Parvaneh Saedi, Jacqueline Schein, Duane Smallis, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.





```
Db 16 GGTCAACGAACTCTTGTAGGCTCTGCCATGACTGAGAGTGGGAGCTTCAC 75
QY 852 GGGGAGCAGCAGCAAGTCTACTCTTCTTCAGGGAGCGGGCAGTGGATCCGACTGCTA 911
Db 76 GGGGAGCAGCAGCAAGTCTACTCTTCTTCAGGGAGCGGGCAGTGGATCCGACTGCTA 135
QY 912 TCCCGACGAGTGGTGGCTGTGTGGCCCGTGTCTGCAGAGGGCGATATGGGGCGCAG 971
Db 136 TCCCGAGCAGTGTGNTGCTGTGTGGCCCGTGTCTGCAGAGGGCGATATGGGGCGCAG 195
QY 972 GACCTGAGAGAGTGGACCACTTCTGAAGGGCGGGCTGGCATGCTTGGCCGAA 1031
Db 196 GACCTGAGAGAGTGGACCACTTCTGAAGGGCGGGCTGGCATGCTTGGCCGAA 255
QY 1032 CTGGCAGCTCTACTTCAACCACTGAGCGATGCAACCTGCAAGACACTCTTGCA 1091
Db 256 CTGGCAGCTCTACTTCAACCACTGAGCGATGCAACCTGCAAGACACTCTTGCA 315
QY 1092 CAACACCACTCTTGGGGTTTTTCAGCAGAGTGGGGTGACATGTACCTGTGGCAT 1151
Db 316 CAACACCACTCTTGGGGTTTTTCAGCAGAGTGGGGTGACATGTACCTGTGGCAT 375
QY 1152 CTGTAGTACCACTTGGAGAGATCCAGCGGGTGTGGAGGGCCCTATAGGAGTACCA 1211
Db 376 CTGTAGTACCACTTGGAGAGATCCAGCGGGTGTGGAGGGCCCTATAGGAGTACCA 435
QY 1212 TGAGGAGCCCAAGAGTGGGACCGCTACACTGACCCTGACCACTCTGGGCTGGCTC 1271
Db 436 TGAGGAGCCCAAGAGTGGGACCGCTACACTGACCCTGACCACTCTGGGCTGGCTC 495
QY 1272 GTGCATTAAACAAGTGGCATCGCGCCACGGCTACACCACTCTCTGGAGCTACCCGACA 1331
Db 496 GTGCATTAAACAAGTGGCATCGCGCCACGGCTACACCACTCTCTGGAGCTACCCGACA 555
QY 1332 CATCCCAACTTCTGTCAAGAACACCCCGTGTATGGAGAGCAGTGGGGCTTCGGTGAG 1391
Db 556 CATCCCAACTTCTGTCAAGAACACCCCGTGTATGGAGAGCAGTGGGGCTTCGGTGAG 615
QY 1392 CCGCCCGCTGTCTGTGAGAGAGGGCACAATTCAACCCACTGTGGGCGGCGGTTAC 1451
Db 616 CCGCCCGCTGTCTGTGAGAGAGGGCACAATTCAACCCACTGTGGGCGGCGGTTAC 675
QY 1452 AGGACTTGATGAGCACCCTATACAGTGTCTTATTTGGCACAGAGACGGCTGGCTGCT 1511
Db 676 AGGACTTGATGAGCACCCTATACAGTGTCTTATTTGGCACAGAGACGGCTGGCTGCT 735
QY 1512 CAAGGCTGTAGCCTGGGGCGCGTGGGTACCTGATGAGAGCTGCACAGCTGTTGACCA 1571
Db 736 CAAGGCTGTAGCCTGGGGCGCGTGGGTACCTGATGAGAGCTGCACAGCTGTTGACCA 795
QY 1572 GGAGCCCATGAGAGGCTGTGTCTATCTGAGAGCAAGAGTGTCTTTGGCGGCTCCG 1631
Db 796 GGAGCCCATGAGAGGCTGTGTCTATCTGAGAGCAAGAGTGTCTTTGGCGGCTCCG 855
QY 1632 CTCTCAGCTGTGTGACAGTCCCGGTGGCCGACATGACATAAAGTATCGCTCTGTGACAGTGT 1691
Db 856 CTCTCAGCTGTGTGACAGTCCCGGTGGCCGACATGACATAAAGTATCGCTCTGTGACAGTGT 915
QY 1692 TGTCCGCGCGGGAGCCCTATTTGGCGCTGAGAGCGTCAACACAGCCGCTGTGGCCGT 1751
Db 916 TGTCCGCGCGGGAGCCCTATTTGGCGCTGAGAGCGTCAACACAGCCGCTGTGGCCGT 975
QY 1752 GGGTGGCCACTTTGGATCTTTACTGATCCAGCATGTGATGACCTGAGACTTCAGGCA 1811
Db 976 GGGTGGCCACTTTGGATCTTTACTGATCCAGCATGTGATGACCTGAGACTTCAGGCA 1035
QY 1812 TTGCAACCTCCGTGGCAGTAAAGTACAGTCAAGGCCCACTTCCAAAAACATCAGGTGT 1871
Db 1036 CTGCAACCTCCGTGGCAGTAAAGTACAGTCAAGGCCCACTTCCAAAAACATCAGGTGT 1092
QY 1872 GGGGGGACAGACCTGTGTCTCCCTGGCAGCTCTCTCAACTTGGCCCTGCCGACTTC 1931
```

```
Db 1093 GGGGGGACAGACCTGTGTCTCCCTGCCACCTCTCTCAACTTGGCCCATGCCCGCTG 1152
QY 1932 CAACCCGAGG 1942
Db 1153 GACCTTTGGGG 1163
```

Search completed: August 19, 2003, 14:58:53  
Job time : 5567 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:45:35 ; Search time 43 Seconds

(without alignments)  
1332.943 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219

Sequence: 1 MAPHNAVMLAARLMLGLIG.....PCHLSSNLALPDSPNESSV 596

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parama

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2720	84.5	834	2	M-sema F protein p
2	1149	35.7	782	2	semaphorin C - mou
3	992.5	30.8	760	2	semaphorin B - mou
4	934.5	29.0	772	2	collapsin - chicke
5	932.5	29.0	771	2	semaphorin III pre
6	931.5	28.9	772	2	semaphorin D - mou
7	892	27.7	749	2	semaphorin V - hum
8	872.5	27.1	748	2	semaphorin A - mou
9	858	26.7	666	2	semaphorin III - m
10	842	26.2	753	2	semaphorin III fam
11	774	24.0	751	2	semaphorin E - mou
12	739	23.0	712	2	hypothetical prote
13	729.5	22.7	1074	2	semaphorin F preu
14	708.5	22.0	730	2	semaphorin IV preu
15	685	21.3	724	2	semaphorin II prec
16	658	20.4	656	2	semaphorin I - fru
17	597	18.5	711	2	semaphorin I - fru
18	406.5	12.6	653	2	semaphorin I preu
19	368	11.4	676	2	hypothetical prote
20	236.5	7.3	1945	2	plexin A - fruit f
21	213	6.6	1894	2	plexin A - precursor
22	207	6.4	1905	2	plexin 1 precursor
23	198	6.2	1872	2	plexin 3 precursor
24	196.5	6.1	1884	2	plexin 2 precursor
25	186	5.8	403	2	A39R protein - vac
26	183	5.7	441	2	hypothetical prote
27	166.5	5.2	1568	2	semaphorin recepto
28	138.5	4.3	2051	2	plexin B - fruit f
29	118.5	3.7	620	2	hypothetical prote

102(b)  
c3b, a

30	116.5	3.6	1375	1	JCS148	hepatocyte growth
31	116	3.6	1425	2	T30811	hepatocyte growth
32	115	3.6	295	2	J01775	saliv protein - v
33	114.5	3.6	1806	2	T23298	hypothetical prote
34	108.5	3.4	1047	2	B71402	hypothetical prote
35	107	3.3	1379	1	S01254	hepatocyte growth
36	102	3.2	577	2	A37779	histocompatibility
37	99.5	3.1	860	2	F71000	hypothetical prote
38	98.5	3.1	1154	2	S39536	paraspinal crystal
39	97.5	3.0	523	2	T04825	hypothetical prote
40	96	3.0	552	2	AD1824	periplasmic oligop
41	96	3.0	1503	2	T18266	cycloindulo-oligos
42	95	3.0	1926	2	S01169	beta-glucosidase c
43	94	2.9	1428	2	C85079	hypothetical prote
44	94	2.9	1607	2	T03022	MAP kinase kinase
45	93.5	2.9	684	1	HJECDA	helicase (EC 3.6.1

## ALIGNMENTS

## RESULT 1

S66498 M-sema F protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 24-Nov-1999

C:Accession: S66498

R:Inagaki, S.; Furuyama, T.; Iwahashi, Y.

FEBS Lett. 370, 269-272, 1995

A:Title: Identification of a member of mouse semaphorin family.

A:Reference number: S66498; MID:95385809; PMID:7656991

A:Accession: S66498

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-834 <IN>

A:Cross-References: EMBL:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599

C:Superfamily: semaphorin

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-834/Product: M-sema F protein #status predicted <MAT>

Query Match	Score	Length	DB 2;	Length	834;
Best Local Similarity	84.5%	83.1%	Pred. No. 1.3e-216;		
Matches 501; Conservative	41;	Mismatches	47;	Indels	14; Gaps 4;
OY	1	MAPHNAVMLAARLMLGLIGAEVMMNLVPRKTVSSGELATVVRFSQGTGIDPFLTLE	60		
DB	1	MAPHNAVMLAARLMLGLIGAEVMMNLVPRKTVSSGELATVVRFSQGTGIDPFLTLE	60		
OY	61	PTGLLYVGAREALFAFSMEALBELQGAISWEAVEKTECIORGNQTECFNIRELPY	120		
DB	61	HSGLLYVGAREALFAFSMEALBELQGAISWEAVEKTECIORGNQTECFNIRELPY	120		
OY	121	NASHLYVCGTVAFOKRCYVNNLTFTLEGEFEDGKPKYPAPAGHGLVDGELYSAT	180		
DB	121	NASHLYVCGTVAFOKRCYVNNLTFTLEGEFEDGKPKYPAPAGHGLVDGELYSAT	180		
OY	121	NSSHLYVCGTVAFOKRCYVNNLTFTLEGEFEDGKPKYPAPAGHGLVDGELYSAT	180		
DB	121	NSSHLYVCGTVAFOKRCYVNNLTFTLEGEFEDGKPKYPAPAGHGLVDGELYSAT	180		
OY	181	LNFLGTEPIILRNKGPMSKTEYLAFVLPNPFHFGASVYVESGFTGDDKRYEPR	240		
DB	181	LNFLGTEPIILRNKGPMSKTEYLAFVLPNPFHFGASVYVESGFTGDDKRYEPR	240		
OY	181	LNFLGTEPIILRNKGPMSKTEYLAFVLPNPFHFGASVYVESGFTGDDKRYEPR	240		
DB	181	LNFLGTEPIILRNKGPMSKTEYLAFVLPNPFHFGASVYVESGFTGDDKRYEPR	240		
OY	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARIAGSAPNMOLYFNOLAM	300		
DB	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARIAGSAPNMOLYFNOLAM	300		
OY	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARIAGSAPNMOLYFNOLAM	300		
DB	241	ERAVESDCYAEQVAVARVAVRCGDMGARTLQKRTTFLKARIAGSAPNMOLYFNOLAM	300		
OY	301	HTLODTSWNTTFFGVFOQKQMDYLSAICEYOLEIQVFGPPTKEIEEOKMDRYD	360		
DB	301	HTLODTSWNTTFFGVFOQKQMDYLSAICEYOLEIQVFGPPTKEIEEOKMDRYD	360		
OY	301	HTLODTSWNTTFFGVFOQKQMDYLSAICEYOLEIQVFGPPTKEIEEOKMDRYD	360		
DB	301	HTLODTSWNTTFFGVFOQKQMDYLSAICEYOLEIQVFGPPTKEIEEOKMDRYD	360		
OY	361	PVPSRPSCINNNRRHGGTSSLELDPIILNFVKKHPLMEQVOPRMSRPLVYKGTNE	420		
DB	361	PVPSRPSCINNNRRHGGTSSLELDPIILNFVKKHPLMEQVOPRMSRPLVYKGTNE	420		
OY	361	PVPSRPSCINNNRRHGGTSSLELDPIILNFVKKHPLMEQVOPRMSRPLVYKGTNE	420		
DB	361	PVPSRPSCINNNRRHGGTSSLELDPIILNFVKKHPLMEQVOPRMSRPLVYKGTNE	420		
OY	421	THLVADRYTGIDGATYVLFITGTGDMGLKAVSLGPMVHLIBELQLFDOEPARSLVLSOS	480		
DB	421	THLVADRYTGIDGATYVLFITGTGDMGLKAVSLGPMVHLIBELQLFDOEPARSLVLSOS	480		



collapse - chicken  
 C:Species: Gallus gallus (chicken)  
 C>Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 24-Sep-1999  
 C:Accession: A49069  
 R: Luo, Y.; Raper, J.A.  
 Cell 75, 217-227, 1993  
 A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neuron  
 A:Reference number: A49069; MUID:94006554; PMID:8402908  
 A:Accession: A49069  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-772 <LUD>  
 A:Cross-references: GB:002528; NID:9410078; PID: AACS9638.1; PID:9410079  
 C:Superfamily: semaphorin

Query Match 29.0%; Score 934.5; DB 2; Length 772;  
 Best Local Similarity 38.2%; Pred. No. 6,2e-69;  
 Matches 207; Conservative 93; Mismatches 203; Indels 37; Gaps 14;

QY 8 WLALRLWGLG-----IGAEVWNNLVPRKTVSSGELA--TVARRFSOTGIODELTLTL 58  
 DB 3 WLRGIALSLGLVLLAGRNCOHKNVPRKLSYKEMLENNVTFNGLANSSVHTFL 62

QY 59 TEPTGLLVGAREALFAFMSMELELQALISWAPVEKTECIQCKNNOTECFNTRFLQ 118  
 DB 63 DEERSRLVYGAKDIFSFNLVNIKEYOKIWPVSHSRDECKWAGKDLRCCANFIKYLK 122

QY 119 PYNASHLYVCGTYAFQPKCTVYNNLT-----FTLEHGFEDGKCKPYDPAKGAGLLV 172  
 DB 123 TYNTHLYACGTGAFHPICTYIEVGSHPEDNIFKEDSHFENGKRSKSPYDKLLTASLLY 182

QY 173 DGLYSATLNNFLGTEPIILRNMGPHHSKTE-VIAFWLNEPHVGSAYVPSVSGFTGD 231  
 DB 183 DGLYSATLADFMGRDFAIFRTLGHHHPIRTEQHSRWLNDPRFISAHLSNDP---E 239

QY 232 DDKYFFPREAREVSDCAEOVAVARVAVCKGDMGATLQKRTTFILKARLACSP--- 288  
 DB 240 DDKYFFPREAREVSDCAEOVAVARVAVCKGDMGATLQKRTTFILKARLACSP--- 288

QY 289 NMOLYFNOLQAMHTLQDTSWHTTFFGVFOAQMGDMYLSAICEYOLEEIQRYFEGPYREY 348  
 DB 300 GIDHFDLQDVFLNMSKDPKNPIYGVFTTSSNIFKSAVCMYSMDVRRVFLGPRYHR 359

QY 349 HEBAOKMDRYDPPSPRPGSCINNMHRHGYTSSLELPDNLINLVKHPIMEEOVGPRW 408  
 DB 360 DGPYQWVPYQGRVYPRPGTCSKTF--GGFDSYKLDLPDDEYTFARSHPRAYNVFPIIN 417

QY 409 SRPLLVKGTN--FTHLVADRVTLGDCATYTLVFTGDSGLKRAVSL--GPVH----L 460  
 DB 418 SRPLLVKGTN--FTHLVADRVTLGDCATYTLVFTGDSGLKRAVSL--GPVH----L 460

QY 461 IEELQDFDQEP--MRSVLVLSQSKLLFAGSRQSLVQLPVAQCIRY--RSCADCVLARDPYC 517  
 DB 477 LEEMTVF-REPTVISAMKISIRKQQLYIGSATVGSQPLRHRCDDYGRKACACCCCLARDPYC 535

QY 518 AW 519  
 DB 536 AW 537

RESULT 5  
 D49423  
 semaphorin III precursor - human  
 C:Species: Homo sapiens (man)  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-Sep-1999  
 C:Accession: D49423  
 R: Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.  
 Cell 75, 1389-1399, 1993  
 A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone  
 A:Reference number: A49423; MUID:94094332; PMID:8269517  
 A:Accession: D49423  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA

A:Residues: 1-771 <KOI>  
 A:Cross-references: GB:L6081; NID:9799328; PID:AAA65938.1; PID:9436560  
 C:Genetics:  
 A:Gene: SEMA1  
 C:Cross-references: GB:283448  
 C:Superfamily: semaphorin

Query Match 29.0%; Score 932.5; DB 2; Length 771;  
 Best Local Similarity 38.4%; Pred. No. 9e-69;  
 Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;

QY 8 WL--LAARLWGLGAEVW---NLVPRKTVSSGELATVARRFSOTGI--QDFLTLTL 58  
 DB 3 WLRVLCFLWGLVLLARANYQNGKNVPRKLSYKEMLENNVTFNGLANSSVHTFL 62

QY 59 TEPTGLLVGAREALFAFMSMELELQALISWAPVEKTECIQCKNNOTECFNTRFLQ 118  
 DB 63 DEERSRLVYGAKDIFSFNLVNIKEYOKIWPVSHSRDECKWAGKDLRCCANFIKYLK 122

QY 119 PYNASHLYVCGTYAFQPKCTVYNNLT-----FTLEHGFEDGKCKPYDPAKGAGLLV 172  
 DB 123 AYNGHLYACGTGAFHPICTYIEIGHPEEDNIFKLENSHFENGKRSKSPYDKLLTASLLI 182

QY 173 DGLYSATLNNFLGTEPIILRNMGPHHSKTE-VIAFWLNEPHVGSAYVPSVSGFTGD 231  
 DB 183 DGLYSATLADFMGRDFAIFRTLGHHHPIRTEQHSRWLNDPRFISAHLSNDP---E 239

QY 232 DDKYFFPREAREVSDCAEOVAVARVAVCKGDMGATLQKRTTFILKARLACSP--- 288  
 DB 240 DDKYFFPREAREVSDCAEOVAVARVAVCKGDMGATLQKRTTFILKARLACSP--- 288

QY 289 NMOLYFNOLQAMHTLQDTSWHTTFFGVFOAQMGDMYLSAICEYOLEEIQRYFEGPYREY 348  
 DB 300 GIDHFDLQDVFLNMSKDPKNPIYGVFTTSSNIFKSAVCMYSMDVRRVFLGPRYHR 359

QY 349 HEBAOKMDRYDPPSPRPGSCINNMHRHGYTSSLELPDNLINLVKHPIMEEOVGPRW 408  
 DB 360 DGPYQWVPYQGRVYPRPGTCSKTF--GGFDSYKLDLPDDEYTFARSHPRAYNVFPIIN 417

QY 409 SRPLLVKGTN--FTHLVADRVTLGDCATYTLVFTGDSGLKRAVSL--GPVH----L 460  
 DB 418 NRPYIKFDVYQFQIYVDRDAEDG-QYDVMFIDGVYLVKAVSIPKETWHELEVYL 476

QY 461 IEELQDFDQEP--MRSVLVLSQSKLLFAGSRQSLVQLPVAQCIRY--RSCADCVLARDPYC 517  
 DB 477 LEEMTVF-REPTVISAMELSTKQQLYIGSTAGVQPLRHRCDDYGRKACACCCCLARDPYC 535

QY 518 AW 519  
 DB 536 AW 537

RESULT 6  
 I48747  
 semaphorin D - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999  
 C:Accession: I48747  
 R: Buschell, A.W.; Adams, R.H.; Betz, H.  
 Neuron 14, 941-948, 1995  
 A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creat  
 A:Reference number: I48747; MUID:95267431; PMID:7748561  
 A:Accession: I48747  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-772 <RES>  
 A:Cross-references: EMBL:X85993; NID:9854329; PID:CAAS9985.1; PID:9854330  
 C:Genetics:  
 A:Gene: semd  
 C:Superfamily: semaphorin

Query Match 28.9%; Score 931.5; DB 2; Length 772;  
 Best Local Similarity 38.6%; Pred. No. 1.1e-68;









```

0Y 115 FLQOPYNASHLYXGYTAPQPCRYVNMFLFTLEGEEDGCKPYDPAKAGLILYD 173
Db 119 RULL-VGDRLEFGQNAFTPVCTNRNSLNAIEHDOT-SGARCPDYSPOHNSTALLTAG 176
0Y 174 GELYSATLNNFLTETEPILLRNMGPHHSMT-EYLAFLMNEPFIHVSAYPEVSGETGDD 232
Db 177 GELYATAMDFPGRDPAIYRSLIGLPLRLTAQYNSKMLNEPFIHVS---YDIGNT--- 229
0Y 233 DXYFEFFRERAVESDCAEOYVAVARVCKDMDGASGLTQKRTFTFLKRLACAP-NMQ 291
Db 230 ---YFFERENAIVBHD-CKGVFSRAARCKMDIGRFLLEDWTETEMKRLKNSRGEVY 285
0Y 292 LYFNQOAMHTLDDTSMHNTTFPGFOAQMGDMYLSACEYDLEEFORFEBGPKYHEE 351
Db 286 FYFNELOSTFLEELD---LYIGIFTTNVAISANAVCFWLSAIQAQFSGPFKYQENS 341
0Y 352 AOKMDRYTDPVSPSPGSCINNMRHRRGYTSS---LELPDNLNFIYKKAHPLMEBOYCPW 408
Db 342 RSAM---LPYPRNP-----HFQCGTVQDGLVNLTERMLDQAQKFLYVHEVQOPVT 390
0Y 409 SRPLATYKKGNFHLVADRYTGLDGATYTVLEFIGNDQWLKA-VSLGFWYH--LLEELD 465
Db 391 TVPSFEDNRSRFEHAVADVQGRALVH-IYILAVDYCTIKKVRVPLNQTSSSCLLEIE 449
0Y 466 LEFDQ---EPMRSLVLSQSKRLFLFAGRSQALVOLPYADCIKYSQADCVIARPDYCAMSVN 522
Db 450 LEPERREREPRIHQILHSQSVLFGVLRHEHYVRIPLKRCQFYFTRSTICGADPDYCGMDVY 509
0Y 523 TSRC-----VAVCGHRS 534
Db 510 MKKCVSLESLMTQWEOGISACPTRNLTVDGHEG 544

```

[illegible]

RESULT 14  
JH0798  
fasciclin IV precursor - American bird grasshopper  
C:Species: Schistocerca americana (American bird grasshopper)  
C:Date: 30-Sep-1993 #sequence\_rev150 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: JH0798  
R:Kolodkin, A.; Mathes, D.J.; O'Connor, T.P.; Patel, N.H.; Admon, A.; Bentley, D.; Goode,  
Neuron 9, 831-845, 1992  
A:Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance in  
A:Reference number: JH0798; MUID:93040225; PMID:141898  
A:Accession: JH0798  
A:Molecule type: mRNA  
A:Residues: 1-730 <KOL>  
A:Cross-References: GB:I.00709; NID:g160844; PID:g160845  
A:Experimental source: embryo  
C:Comment: This protein plays a role in growth cone guidance in the developing central  
C:Keywords: glycoprotein; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-730/Product: fasciclin IV #status predicted <MGT>  
F:23-627/Domain: extracellular #status predicted <EXT>  
F:628-653/Domain: transmembrane #status predicted <TM>  
F:653-730/Domain: intracellular #status predicted <INT>  
F:44,71,63,267,360,535/Binding site: carbohydrate (Asn) (covalent) #status predicted

RESULT 15  
 C49423  
 semaphorin II precursor - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 07-May-1999  
 C:Accession: C49423  
 R:Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.  
 Cell 75, 1369-1399, 1993  
 A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth co  
 A:Reference number: A49423; MUID:94094332; PMID:826517  
 A:Accession: C49423  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
 A:Molecule type: mRNA  
 A:Residues: 1-724 <KOL>  
 A:Cross-references: GB:I26083  
 C:Genetics:  
 A:Gene: sema II  
 A:Cross-references: FLYBase:FBgn0011260  
 C:Superfamily: semaphorin  
  
 Query Match 21.3%, Score 685, DB 2, Length 724;  
 Best Local Similarity 32.2%, Pred. NO, 2.5e-48;  
 Matches 185; Conservative 92; Mismatches 219; Indels 78; Gaps 23;

Query Match	22.0%	Score 708.5	DB 2	Length 730
Best Local Similarity	31.3%	Pred. No. 2.8e-50		
Matches 190	Conservative 107	Mismatches 245	Indels 65	Gaps 21
OY	6	AVMLLAARLMTGLGAEVMMNLV-PRKTVSSGELATVVRFRSOTGIDFLTTLTPTGL	64	
DB	4	ALVAVALALLM-VALHAAAMVNDVSPKMTVDPGE-ERVRFPLGNESKIDHFKLLERDHNLS	61	
OY	65	LYVGAREALFAFSMEALE--LQCAISWEAPYEKRTKCIQKGNNOGTECFNFTFLDPYNA	122	
DB	62	L-VGARNIVYINSLRDLETFEQRLEIEMHSGAHRELCTYKGR-SEDDCQNYIRVLAKIDD	119	
OY	123	SHLYVCGTAFQKCYIVNMILTFLEHGEF-----EDCKGKPPYDPAKSHAGILVDGELY	177	
DB	120	DWYLCTGTAAVKLCRR-----VALKDGDYVEKEKRGCLCFPPDHNSTAIFYSGOLY	174	

```

QY 42 VARESOTGIQDPLTLTLTEPTGLLYVGAEEALFAESMEALLE---QGAISWEAPYEKKT 97
DB 58 VEEFN-CGLLYRTFTEPMNEDRDTLLXYGAMDREVRVLIQAISSNCNRDAINLEPTRDDVY 116
QY 98 ECIQKGNQGTCECFNFIREFLOPYN-ASHLYVCGTAFORCKCYVMNLTTT-LEHGFEEOG 155
DB 117 SCVSKGKSQIFQCKKNHVRVLIQSDQDGRLYVCGTGNHNKR-DVYIYANTHLPRSEYVIG 175
QY 156 KG-----KCPYDPAKGAGLAVDG-----ELYSATLNNFLGTEPIITLRMGPHSKM--- 202
DB 176 VGLGIAKCYDPLDNSTAIYVENGNRGGLPGLXSGINAEFTKADYIYIFRDLTYNTSAKRL 235
QY 203 -----TEYAFWLNIEPHFVGSAYAVESVGSFTGDDRYVEFFERAVESDCYAOYYA 255
DB 236 EYKFRKTLAYDKMDLKNPVEVGSF-----DIGEV-----VYFEERETAVEYIMGCAVS 285

```

```
QY 256 RVARVCKGMGARTLQRMWTFPLKRLACS-APNQLYENQLQAMHTLQDTSMHNTTF 314
Db 286 RVARVCKKDVGGKMLAHMMATYLLKARLNCISGEFPFYENEIQSYQLPS---DKSRFF 342
QY 315 GVFOAQMGDMYLSAICEYQLFEIQRFEGPYKEHYHEAQKW-DRYTDPVPSPPRGSCINN 373
Db 343 ATFTTSTNGLISAVCSFHINEIQAAFNGKFKQSSNSAMLPVLNSRVPEPRPGTCVND 402
QY 374 WHRRHGYTSSLELPDNLNFVKRHPLMEDQVGRWRSPRLVYKGTNFTHLVADRV-TGLD 432
Db 403 -----TSN--LPDTVLNFRSHPLMDKAVNHEHNNPYKKRDLYFTKLVDKIRIDIL 453
QY 433 GATYTYVLFCTGDCMLLKAIV-----SLGPNVHLIEQLDQEPMSLVSQSKLLP 485
Db 454 NOEYIYYVGTNGRIYKIYOYRNGESLS---KLIDIEVAPEAIAQYMEISQTRKSLY 510
QY 486 AGSRSQLVOLPYADC-IKYRSCADCYLADPYCAMSVNTSRCAVAGHGSLLIQHVMTS 544
Db 511 IGTDRIKOIDLAMCNRRTDNCRCV--RDPYCGMDKEANTC---RPELDLLODV-AN 563
QY 545 DTSGICNLGRSKIOSGPPXPKNITVAVGTDVLVLP 578
Db 564 ETSIDIC-----SSVLKKKIYVTFYQGSVHLGC 590
```

Search completed: August 7, 2003, 13:50:46  
Job time : 46 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:41:55 ; Search time 83 Seconds

(without alignments)  
1139.772 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219  
Sequence: 1 MAPRAWMLAARLGLGIC.....PCHLSNLAIPDSNPBSSV 596

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3217	99.9	596 22 AAB48374	Human SEC7 protein
2	3115.5	96.8	590 22 AAB48373	Human SEC6 protein
3	3113.5	96.7	833 22 AAE03640	Human extracellular
4	3107.5	96.5	624 22 AAB48378	Human SEC11 protei
5	3089.5	96.0	833 23 AAE18213	Human MO15a protei
6	3085.5	94.9	833 23 AAE18214	Human MO15c protei
7	3055.5	94.9	833 23 AAE18215	Human MO15c protei
8	2927.5	90.9	805 23 ABB98402	Human NOV2, a se
9	2918	90.6	886 23 ABB97963	Human protein sequ

10	2720	84.5	834 22 AAB61238	Murine M-Sema-F pr
11	2635	81.9	785 23 ABU60951	Lung specific prot
12	2635	81.9	785 23 ABU18435	Breast specific re
13	1685.5	52.4	569 23 ABB97213	Novel human protei
14	1240	38.5	832 22 AAE03818	Human gene 1 encod
15	1240	38.5	832 22 ABB64522	Human albumin fusi
16	1240	38.5	837 21 AAY99410	Human PRO1480 (UNO
17	1240	38.5	837 22 AAU29250	Human PRO polypept
18	1240	38.5	837 23 ABB65159	Protein of the inv
19	1240	38.5	837 23 ABB65100	Hypoxia-induced pr
20	1240	38.5	837 24 ABR48242	Human bladder can
21	1240	38.5	837 24 ABU71338	Human PRO1480 prot
22	1240	38.5	837 24 ABU65795	Human secreted/Lra
23	1240	38.5	837 24 ABU66128	Novel human secret
24	1240	38.5	837 24 ABU67632	Human secreted/Lra
25	1240	38.5	837 24 ABU65490	Human PRO polypept
26	1240	38.5	837 24 ABU65626	Human PRO polypept
27	1240	38.5	837 24 ABU56162	Human secreted/Lra
28	1240	38.5	837 24 ABU56640	Human secreted/Lra
29	1240	38.5	837 24 ABU57157	Human PRO polypept
30	1240	38.5	837 24 ABU10736	Human secreted/Lra
31	1240	38.5	837 23 ABB97964	Human protein sequ
32	1226	38.1	843 23 AAG68301	Human semaphorin G
33	1226	38.1	870 23 AAG68303	Human semaphorin G
34	1225.5	38.1	838 23 AAG68302	Human semaphorin G
35	1225.5	38.1	861 18 AAM17658	Human CD100 antige
36	1225.5	38.1	865 23 AAG68304	Human semaphorin G
37	1222.5	38.0	861 19 AAW58540	Human semaphorin
38	1222.5	38.0	861 22 AAB81035	Murine CD100 amino
39	1222.5	38.0	861 22 AAB81251	Mouse CD100 protei
40	1221	37.9	245 22 AAB20160	Human protein SECP
41	1219.5	37.9	862 18 AAM17657	Human CD100 antige
42	1219.5	37.9	862 22 AAB81036	Human CD100 amino
43	1219.5	37.9	862 22 AAB51252	Human CD100 protei
44	1219	37.9	791 23 AAU77413	Human NOV7 protein
45	1182.5	36.7	573 22 AAM06786	Human foetal prote

#### ALIGNMENTS

RESULT 1	
AAB48374	standard; Protein; 596 AA.
ID	AAB48374;
AC	20-APR-2001 (first entry)
XX	Human SEC7 protein sequence (clone ID 20422974.2);
DE	SEC7: cytosolic; gynecological; gene therapy; screening assay; human;
KW	SEC7; chromosomal mapping; forensic biology; cell proliferation; cancer;
KM	cell differentiation; immune associated disorder; gestational disease.
XX	
OS	Homo sapiens.
XX	
XX	
FT	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	/note="signal peptide"
FT	21..596
FT	/note="mature protein"
XX	
PN	MO200078802-A2.
XX	
PD	28-DEC-2000.
XX	
PF	23-JUN-2000; 2000WO-US17328.
XX	
PR	23-JUN-1999; 9905-0140584.
PR	20-JUL-1999; 9905-0144722.
PR	16-SEP-1999; 9905-0154520.
PR	22-JUN-2000; 2000US-0604286.

XX (CURA-) CURAGEN CORP.  
 PA Shinketsu RA, Fernandes E, Vernet C, Yang M, Boldog FL;  
 PI Hermann JL;  
 XX WPI; 2001-071385/08.  
 DR N-PSDB; AAC84888.  
 XX  
 XX Polynucleotides encoding SECX proteins useful for treating disease  
 PT characterized by an aberrant level of cell proliferation and/or  
 PT differentiation like cancer or immune associated disorders -  
 Claim 1; Fig 7; 132pp; English.  
 XX  
 XX The invention relates to human SECX polypeptides and polynucleotides  
 CC encoding them. The SECX polypeptides can be expressed by standard  
 CC recombinant methodology. The SECX polypeptides are useful for treating  
 CC or preventing a SECX-associated disorder. The invention is useful in  
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and  
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,  
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and  
 CC methods of treatment (e.g. therapeutic and prophylactic), especially  
 CC disorders characterized by aberrant cell proliferation and/or  
 CC differentiation like cancer or immune associated disorders or gestational  
 CC disease. The present sequence represents a SEC7 protein.  
 CC  
 XX  
 XX Sequence 596 AA:  
 SQ  
 Query Match 99.9%; Score 3217; DB 22; Length 596;  
 Best Local Similarity 99.8%; Pred. No. 3.7e-317;  
 Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 MAPHNAVWMLAARLMGIGIGAEYWNLVPRKTVSSGELATVVRFSQTGIDFLTITL 60  
 1 MAPHNAVWMLAARLMGIGIGAEYWNLVPRKTVSSGELATVVRFSQTGIDFLTITL 60  
 61 PPGGLTVGARREALFAFSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQ 120  
 61 PPGGLTVGARREALFAFSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQ 120  
 121 NASHLVGCGTVAPORPCCTVYNNLFTLEGEFEDGKCKGPRDPAKHAGLVDGELYSAT 180  
 121 NASHLVGCGTVAPORPCCTVYNNLFTLEGEFEDGKCKGPRDPAKHAGLVDGELYSAT 180  
 181 LNPFGLTEPILIRNMGPHHSMKTEYLAFWLNPEHFVSAYVPSVSGFGDDDKVYFFR 240  
 181 LNPFGLTEPILIRNMGPHHSMKTEYLAFWLNPEHFVSAYVPSVSGFGDDDKVYFFR 240  
 181 LNPFGLTEPILIRNMGPHHSMKTEYLAFWLNPEHFVSAYVPSVSGFGDDDKVYFFR 240  
 241 ERAVESDCYAEQVAVARVAVCKGDMGARTLQKKTTPFKARLACAPWMLYFNOLQAM 300  
 241 ERAVESDCYAEQVAVARVAVCKGDMGARTLQKKTTPFKARLACAPWMLYFNOLQAM 300  
 301 HTLDPTSMHNTFFGFGQAMGDMYLSACEYQLEIYQVFEGSPYKEHBEAOKMRYND 360  
 301 HTLDPTSMHNTFFGFGQAMGDMYLSACEYQLEIYQVFEGSPYKEHBEAOKMRYND 360  
 361 PVPSPRPGSCINMHRHRYGYSLELDPNLFVKKHPLMEQVGRBMSRPLLYKKGTFN 420  
 361 PVPSPRPGSCINMHRHRYGYSLELDPNLFVKKHPLMEQVGRBMSRPLLYKKGTFN 420  
 421 THLVADRVTVGLDGAATTVLFIGTGDGWLKAVSLGFWHLIELOLPDOEPMSLVSOS 480  
 421 THLVADRVTVGLDGAATTVLFIGTGDGWLKAVSLGFWHLIELOLPDOEPMSLVSOS 480  
 481 KKLTFAGSRSOLVOLPVADCIKRYSCADCVILADPVCASVMSRCAVAAAGHGSLI 540  
 481 KKLTFAGSRSOLVOLPVADCIKRYSCADCVILADPVCASVMSRCAVAAAGHGSLI 540  
 541 VMTSDTSGICNLGRSKIQSGPYKNTTVAGTDLVPLCHLSSNIALPDSNPRESSV 596  
 541 VMTSDTSGICNLGRSKIQSGPYKNTTVAGTDLVPLCHLSSNIALPDSNPRESSV 596

RESULT 2  
 AAB48373  
 ID AAB48373 standard; Protein; 590 AA.  
 XX  
 XX AAB48373;  
 XX  
 XX 20-APR-2001 (first entry)  
 XX  
 XX Human SEC6 protein sequence (clone ID 20422974.0.133).  
 XX  
 KW SECX; cytosolic; gynecological; gene therapy; screening assay; human;  
 KW SEC6; chromosomal mapping; forensic biology; cell proliferation; cancer;  
 KW cell differentiation; immune associated disorder; gestational disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Peptide 1..20  
 FT Protein /note="signal peptide"  
 FT Protein 21..590  
 FT Protein /note="mature protein"  
 XX  
 PN W0200078802-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 PF 23-JUN-2000; 2000MO-US17328.  
 XX  
 PR 23-JUN-1999; 990S-0140584.  
 PR 20-JUL-1999; 990S-0144722.  
 PR 16-SEP-1999; 990S-0154520.  
 PR 22-JUN-2000; 2000US-0604286.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 PI Shinketsu RA, Fernandes E, Vernet C, Yang M, Boldog FL;  
 PI Hermann JL;  
 DR WPI; 2001-071385/08.  
 DR N-PSDB; AAC84887.  
 XX  
 PS Claim 1; Fig 6; 132pp; English.  
 XX  
 CC The invention relates to human SECX polypeptides and polynucleotides  
 CC encoding them. The SECX polypeptides can be expressed by standard  
 CC recombinant methodology. The SECX polypeptides are useful for treating  
 CC or preventing a SECX-associated disorder. The invention is useful in  
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and  
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,  
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and  
 CC methods of treatment (e.g. therapeutic and prophylactic), especially  
 CC disorders characterized by aberrant cell proliferation and/or  
 CC differentiation like cancer or immune associated disorders or gestational  
 CC disease. The present sequence represents a SEC6 protein.  
 CC  
 XX  
 XX Sequence 590 AA:  
 SQ  
 Query Match 96.8%; Score 3115.5; DB 22; Length 590;  
 Best Local Similarity 99.1%; Pred. No. 7.5e-307;  
 Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 1 MAPHNAVWMLAARLMGIGIGAEYWNLVPRKTVSSGELATVVRFSQTGIDFLTITL 60  
 1 MAPHNAVWMLAARLMGIGIGAEYWNLVPRKTVSSGELATVVRFSQTGIDFLTITL 60  
 61 PPGGLTVGARREALFAFSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQ 120  
 61 PPGGLTVGARREALFAFSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQ 120

```

OY 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAKGAGLVLDGELYSAT 180
DB 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAKGAGLVLDGELYSAT 180
OY 181 LNNFLGTEPIILRNNGPHSHSKTEYLAFWLNBPHEVGSAYVPESSGFTGDDDKYFFPR 240
DB 181 LNNFLGTEPIILRNNGPHSHSKTEYLAFWLNBPHEVGSAYVPESSGFTGDDDKYFFPR 240
OY 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQKKTTEFLKARLACSAFPMOLYFNQLOAM 300
DB 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQKKTTEFLKARLACSAFPMOLYFNQLOAM 300
OY 301 HTLQDTSWNTTFFGVFOAQNGDMYLSAICEYOLEETORVPEGPKYKEYHEEAQMDRYTD 360
DB 301 HTLQDTSWNTTFFGVFOAQNGDMYLSAICEYOLEETORVPEGPKYKEYHEEAQMDRYTD 360
OY 361 PVSPRPGSCINNMHRRHGYSSLELPTNLIIFVKKHPLMEQVGPFRMSRPLLYKKGTFE 420
DB 361 PVSPRPGSCINNMHRRHGYSSLELPTNLIIFVKKHPLMEQVGPFRMSRPLLYKKGTFE 420
OY 421 THLVADRTYGLDGAITYVLFTGTGDMPLKAVSLGPMVHLTELOLFOEPPARSLVLSOS 480
DB 421 THLVADRTYGLDGAITYVLFTGTGDMPLKAVSLGPMVHLTELOLFOEPPARSLVLSOS 480
OY 481 KLLFAGRSQOLVOLPVADCIKYNRCADCVLARDPYCAMSVMTSRCAVAGHFGSLTIOH 540
DB 481 KLLFAGRSQOLVOLPVADCIKYNRCADCVLARDPYCAMSVMTSRCAVAGHFGSLTIOH 540
OY 541 VMTSDTSGICMLRGSKIQSPKNTYVAGNDVLVPLCHLS 582
DB 541 VMTSDTSGICMLRGSKIQSPKNTYVAGNDVLVPLCHLS 582

```

RESULT 3  
AAE03640  
ID AAE03640 standard; Protein: 833 AA.

AAE03640;  
06-AUG-2001 (first entry)

Human extracellular matrix and cell adhesion molecule-4 (XMAD-4).  
Human: extracellular matrix and cell adhesion molecule; XMAD;  
gene therapy; genetic disorder; adrenoleukodystrophy; leukaemia;  
Down's syndrome; cystic fibrosis; Gaucher's disease; myotonic dystrophy;  
sickle cell anaemia; thalassaemia; autoimmune disorder; adenocarcinoma;  
inflammatory disorder; acquired immune deficiency syndrome; AIDS;  
Addison's disease; allergy; anaemia; asthma; atherosclerosis; melanoma;  
Crohn's disease; diabetes mellitus; atopic dermatitis; lymphoma; cancer;  
glomerulonephritis; multiple sclerosis; Grave's disease; osteoarthritis;  
osteoporosis; psoriasis; rheumatoid arthritis; ulcerative colitis;  
infection; cell proliferative disorder; actinic keratosis; myeloma;  
arteriosclerosis; neutropil; anticonvulsant; antithyroid; nephrotropic;  
neuroprotective; dermatological.

Homo sapiens.  
Key Location/Qualifiers  
Peptide 1..23  
Protein /label= Signal\_peptide  
24..833  
/note= "Mature human extracellular matrix and cell  
adhesion molecule (XMAD)"  
53..481  
/note= "Semaphorin domain"  
WO200142285-A2.  
14-JUN-2001.  
05-DEC-2000; 2000WO-US32990.

```

XX 10-DEC-1999; 9905-0172852.
PR 16-DEC-1999; 9905-0172354.
XX
PA (INCY-) INCYTE GENOMICS INC.
PI Yue H, Tang YT, Lai P, Burford N, Azimzai Y, Patterson C;
PI Baughn WR, Lu DM, Shah P, Au-Young J;
DR WPI: 2001-381632/40.
DR N-PSDB: AAD08048.
XX
PT New human extracellular matrix and cell adhesion molecules and
PT polynucleotide sequences encoding them, useful for diagnosis,
PT prevention, treatment of genetic, autoimmune and cell proliferative
PT disorders
XX
PS Claim 1; Page 96-98; 135pp; English.
XX
CC The present sequence is a human extracellular matrix and cell
CC adhesion molecule (XMAD). The XMAD is used for screening a compound for
CC effectiveness as an agonist or antagonist of XMAD. The identified agonist
CC or antagonist are used for treating a disease or condition associated
CC with decreased or increased expression of functional XMAD. The
CC polynucleotides encoding XMAD are useful in somatic or germline gene
CC therapy to correct a genetic deficiency, to express a conditionally
CC lethal gene product and to express a protein which affords protection
CC against intracellular parasites and also for diagnosis of disorders
CC associated with expression of XMAD. They are also used for generating
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequences and to create knock in humanised animals (pigs) or transgenic
CC animals (mice or rats) to model human diseases. Oligonucleotide or longer
CC fragments derived from the polynucleotide sequences may be used as
CC elements on a microarray. Antibodies which specifically bind XMAD may be
CC used for the diagnosis of disorders associated with the expression of
CC XMAD, or in assays to monitor patients being treated with XMAD. Diseases
CC diagnosed, prevented or treated include genetic disorders such as
CC adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's
CC disease, myotonic dystrophy, sickle cell anaemia, thalassaemia,
CC autoimmune/inflammatory disorders such as acquired immune deficiency
CC syndrome (AIDS), Addison's disease, allergies, anaemia, asthma,
CC atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis,
CC glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis,
CC osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis,
CC bacterial, fungal, parasitic, protozoal and helminthic infections and
CC cell proliferative disorders such as actinic keratosis, arteriosclerosis
CC and cancer including breast, bladder, bone marrow, brain and uterus
CC cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.
XX
SQ Sequence 833 AA;
Query Match 96.7%; Score 3113.5; DB 22; Length 833;
Best Local Similarity 96.3%; Pred. No. 2.1e-306;
Matches 579; Conservative 4; Mismatches 7; Indels 11; Gaps 2;
OY 1 MAPENAVMLAARLNGIGIGAEVMMNLVPRKTVSSGELATVYRRSQTGIDPFLTLTTE 60
DB 1 MAPENAVMLAARLNGIGIGAEVMMNLVPRKTVSSGELATVYRRSQTGIDPFLTLTTE 60
OY 61 PTGLIYVGARREALFAPSMALDQAGISREAVVEKTECTQKGNQNTCEPFIPIFLQY 120
DB 61 PTGLIYVGARREALFAPSMALDQAGISREAVVEKTECTQKGNQNTCEPFIPIFLQY 120
OY 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAKGAGLVLDGELYSAT 180
DB 121 NASHLYVCGTAFQPKCTYVNMFLFTLEHGEFEDGKGCPTDPAKGAGLVLDGELYSAT 180
OY 181 LNNFLGTEPIILRNNGPHSHSKTEYLAFWLNBPHEVGSAYVPESSGFTGDDDKYFFPR 240
DB 181 LNNFLGTEPIILRNNGPHSHSKTEYLAFWLNBPHEVGSAYVPESSGFTGDDDKYFFPR 240
OY 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQKKTTEFLKARLACSAFPMOLYFNQLOAM 300
DB 241 ERAVESDCYAEQVAVARAVAVCKGDMGARTLQKKTTEFLKARLACSAFPMOLYFNQLOAM 300

```

Db 241 ERALESDCYAEVAVARAVKCGDMGANTLQKRTTFLKARLACSPAPMQLYFNOLQAM 300  
 QY 301 HTLQDTSMHNTTFFGVFOAQMDMYLSAICEYOLEEIQRFEGPYKEYHEEAQKMDRYTD 360  
 Db 301 HTLQDTSMHNTTFFGVFOAQMDMYLSAICEYOLEEIQRFEGPYKEYHEEAQKMDRYTD 360  
 QY 361 PVPSPRPGSCINNMHRHGGTSSLELPDNLINFKKHPIMEQOVGRMSRPLLKKGTNF 420  
 Db 361 PVPSPRPGSCINNMHRHGGTSSLELPDNLINFKKHPIMEQOVGRMSRPLLKKGTNF 420  
 QY 421 THLVADRVGTGLDAGATYTVLEFIGDGMWLKAVSLGPMVHLIEELOLFDOPMRSYLVSOS 480  
 Db 421 THLVADRVGTGLDAGATYTVLEFIGDGMWLKAVSLGPMVHLIEELOLFDOPMRSYLVSOS 480  
 QY 481 KLLFAGSRSQLVQLFVADCIKRYSCADCVLADPYCAMSVMYSRCAVAGHRSLLIOH 540  
 Db 481 KLLFAGSRSQLVQLFVADCIKRYSCADCVLADPYCAMSVMYSRCAVAGHRSLLIOH 540  
 QY 541 VMTSDTSGICNLGSKGSK - KVRPTPKNITVYAGTDLVLPCHLSNLAHARMTFCGRDLPAQ 599  
 Db 541 VMTSDTSGICNLGSKGSK - KVRPTPKNITVYAGTDLVLPCHLSNLAHARMTFCGRDLPAQ 599  
 QY 591 P 591  
 Db 600 P 600

RESULT 4  
 AAB48378  
 ID AAB48378 standard; protein; 624 AA.

AC AAB48378;  
 DT 20-APR-2001 (first entry)

Human SEC11 protein sequence (clone ID 20422974.0.132-ext2).

SECCX: cytostatic; gynecological; gene therapy; screening assay; human;  
 SEC11: chromosomal mapping; forensic biology; cell proliferation; cancer;  
 cell differentiation; immune associated disorder; gestational disease.

Homo sapiens.

Key Location/Qualifiers  
 Peptide 1..20  
 Protein /note= "signal peptide"  
 21..624  
 /note= "mature protein"

MO200078802-A2.

28-DEC-2000.

23-JUN-2000; 2000MO-US17328.

23-JUN-1999; 99US-0140584.

20-JUL-1999; 99US-0144722.

16-SEP-1999; 99US-0154520.

22-JUN-2000; 2000US-0604286.

(CURA-) CURAGEN CORP.

Shinketsu RA, Fernandes E, Vernet C, Yang M, Boldog FL;

Herrmann JI;

WPI: 2001-071385/08.

N-PSDB: AAC84892.

Polynucleotides encoding SECCX proteins useful for treating disease

PT characterized by an aberrant level of cell proliferation and/or

PT differentiation like cancer or immune associated disorders -

XX The invention relates to human SECCX polypeptides and polynucleotides  
 CC encoding them. The SECCX polypeptides can be expressed by standard  
 CC recombinant methodology. The SECCX polypeptides are useful for treating  
 CC or preventing a SECCX-associated disorder. The invention is useful in  
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and  
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,  
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and  
 CC methods of treatment (e.g. therapeutic and prophylactic), especially  
 CC disorders characterized by aberrant cell proliferation and/or  
 CC differentiation like cancer or immune associated disorders or gestational  
 CC disease. The present sequence represents a SEC11 protein.

SQ Sequence 624 AA;  
 Query Match 96.5%; Score 3107.5; DB 22; Length 624;  
 Best Local Similarity 99.0%; Pred. No. 5.4e-306;  
 Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 MAPHMAVWLLAARLMLGIGAEYMWNLVPRKTVSSGELATVVRFSOTGIQDFLTLLTE 60  
 Db 1 MAPHMAVWLLAARLMLGIGAEYMWNLVPRKTVSSGELATVVRFSOTGIQDFLTLLTE 60  
 QY 61 PTGLLYVGAREALFARSMLELEQAI SWEAPYEKTECIQKKNNOTECFNRIRLOPY 120  
 Db 61 PTGLLYVGAREALFARSMLELEQAI SWEAPYEKTECIQKKNNOTECFNRIRLOPY 120  
 QY 121 NASHLVYCGTYAPOPCTYVNMFLTLEHGEFEDGKCPYDPAKGNAGLLVYGEYSAT 180  
 Db 121 NASHLVYCGTYAPOPCTYVNMFLTLEHGEFEDGKCPYDPAKGNAGLLVYGEYSAT 180  
 QY 181 LNNFLGTEPIILNNMGPBHSKTEYLAFWLNEDHPFVSAYVPSVSGFTGDDDKYVEFFR 240  
 Db 181 LNNFLGTEPIILNNMGPBHSKTEYLAFWLNEDHPFVSAYVPSVSGFTGDDDKYVEFFR 240  
 QY 241 ERAVESDCYAEVAVARAVKCGDMGARTLQKRTTFLKARLACSPAPMQLYFNOLQAM 300  
 Db 241 ERAVESDCYAEVAVARAVKCGDMGARTLQKRTTFLKARLACSPAPMQLYFNOLQAM 300  
 QY 301 HTLQDTSMHNTTFFGVFOAQMDMYLSAICEYOLEEIQRFEGPYKEYHEEAQKMDRYTD 360  
 Db 301 HTLQDTSMHNTTFFGVFOAQMDMYLSAICEYOLEEIQRFEGPYKEYHEEAQKMDRYTD 360  
 QY 361 PVPSPRPGSCINNMHRHGGTSSLELPDNLINFKKHPIMEQOVGRMSRPLLKKGTNF 420  
 Db 361 PVPSPRPGSCINNMHRHGGTSSLELPDNLINFKKHPIMEQOVGRMSRPLLKKGTNF 420  
 QY 421 THLVADRVGTGLDAGATYTVLEFIGDGMWLKAVSLGPMVHLIEELOLFDOPMRSYLVSOS 480  
 Db 421 THLVADRVGTGLDAGATYTVLEFIGDGMWLKAVSLGPMVHLIEELOLFDOPMRSYLVSOS 480  
 QY 481 KLLFAGSRSQLVQLFVADCIKRYSCADCVLADPYCAMSVMYSRCAVAGHRSLLIOH 540  
 Db 481 KLLFAGSRSQLVQLFVADCIKRYSCADCVLADPYCAMSVMYSRCAVAGHRSLLIOH 540  
 QY 541 VMTSDTSGICNLGSKGSK - KVRPTPKNITVYAGTDLVLPCHLS 582  
 Db 541 VMTSDTSGICNLGSKGSK - KVRPTPKNITVYAGTDLVLPCHLS 581

RESULT 5  
 AAE18213  
 ID AAE18213 standard; protein; 833 AA.

AC AAE18213;  
 DT 07-MAY-2002 (first entry)

Human MOL5a protein.

DE Secreted molecule; MOL5a protein; MOLX; cardiomyopathy; atherosclerosis;  
 KW diabetes; chromosomal disorder; aplastic anemia; psoriasis; scarring;  
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;

KM cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;  
 KM immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;  
 KM HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;  
 KM haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;  
 KM Parkinson's disease; Huntington's disease; muscular disease; stress;  
 KM ocular disease; growth disorder; depression; epilepsy; contraceptive;  
 KM vulnery; osteopathic; haemostatic; tranquilliser; antidepressant;  
 KM analgesic; vasotropic; hypotensive; gene therapy; chromosome 2.  
 XX Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 247 /note- "Encoded by gcc"  
 FT  
 XX  
 XX  
 XX  
 PD 24-JAN-2002.  
 XX  
 XX 03-JUL-2001; 2001WO-US21249.  
 XX  
 XX 03-JUL-2000; 2000US-215854P.  
 PR 03-JUL-2000; 2000US-215856P.  
 PR 03-JUL-2000; 2000US-215902P.  
 PR 07-JUL-2000; 2000US-215855P.  
 PR 07-JUL-2000; 2000US-216586P.  
 PR 17-JUL-2000; 2000US-216722P.  
 PR 17-JUL-2000; 2000US-218622P.  
 PR 17-JUL-2000; 2000US-218992P.  
 PR 27-JUL-2000; 2000US-221285P.  
 PR 14-FEB-2001; 2001US-268734P.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Spaderina SK, Tchernev V, Liu X, Shenoy S, Spytek K, zerhusen B;  
 PI Patubajjan M, Taupier RJ, Stellai L, Grose RM, Szekeres ES;  
 PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;  
 PI Padigaru M;  
 XX  
 XX WPI: 2002-155038/20.  
 DR N-PSDB: AAD28947.  
 XX  
 XX  
 PT Nucleic acids encoding secreted polypeptides, designated MOlx  
 PT polypeptides, useful for treating a MOlx-associated disorder, e.g.  
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -  
 XX  
 XX Claim 1; Page 60; 223pp: English.

CC The patent discloses nucleic acid sequences encoding novel secreted  
 CC molecule (MOlx) polypeptides, designated MOlx polypeptides (i.e. a MOlx  
 CC protein where x is an integer from 1 to 8). Sequences of the invention  
 CC are useful for treating or preventing a MOlx-associated disorder in  
 CC humans. They are useful for treating or preventing cardiomyopathy,  
 CC atherosclerosis and disorders related to cell signal processing and  
 CC metabolic pathway modulation. The MOlx antibodies are useful for  
 CC treating or preventing diabetes and disorders related to cell signal  
 CC processing and metabolic pathway modulation. MOlx sequences are useful  
 CC for the treatment or diagnosis of other MOlx-associated disorders, e.g.  
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,  
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral  
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune  
 CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,  
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,  
 CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,  
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC ocular disease, muscular diseases, growth disorders, loss of libido,  
 CC stress, depression, pain and epilepsy. They are useful for preventing  
 CC chemotherapy side effects and as contraceptives. Sequences of the  
 CC invention are also useful for gene therapy. The present sequence  
 CC is human semaphorin 4C-like protein, MO15a. MO15a gene is localised  
 CC on chromosome 2.  
 XX  
 XX Sequence 833 AA;

Query Match 96.0%; Score 3089.5; DB 23; Length 833;  
 Best Local Similarity 95.7%; Pred. No. 5.8e-304;  
 Matches 575; Conservative 4; Mismatches 11; Indels 11; Gaps 2;

QY	1	MAPHAAVLLAARLALMGIGAEVMMNLVPRKTVSSGELATVVRPSQGTGIDFLTLTTE	60
DB	1	MAPHAAVLLAARLALMGIGAEVMMNLVPRKTVSSGELATVVRPSQGTGIDFLTLTTE	60
QY	61	PTGLLYGARRALAFSNALELQAGISWEAPVEKTCIOKRNQTECFNIFRLQPY	120
DB	61	PTGLLYGARRALAFSNALELQAGISWEAPVEKTCIOKRNQTECFNIFRLQPY	120
QY	121	NASHLYVCGTYAFQPKCTYVNNLTFTLEHGEFEDKGRKCPYDPAPKAGHGLVYDGLYSAT	180
DB	121	NASHLYVCGTYAFQPKCTYVNNLTFTLEHGEFEDKGRKCPYDPAPKAGHGLVYDGLYSAT	180
QY	181	LNNFLGTEPILIRNNGPHHSKTEYLAFWLNPHFVGSAYVPSVSGSTGDDKRYFFER	240
DB	181	LNNFLGTEPILIRNNGPHHSKTEYLAFWLNPHFVGSAYVPSVSGSTGDDKRYFFER	240
QY	241	ERAVSDCYAEOYVARAVRCKGDMGARTLQKRTFLKARLACSAANNQLYFNOLOAM	300
DB	241	ERAVSDCYAEOYVARAVRCKGDMGARTLQKRTFLKARLACSAANNQLYFNOLOAM	300
QY	301	HTLQDSHNTTFEGVFOQMGDMYLSAICEYQLEIQRVEGPKYKEHEEAKMDRYTD	360
DB	301	HTLQDSHNTTFEGVFOQMGDMYLSAICEYQLEIQRVEGPKYKEHEEAKMDRYTD	360
QY	361	PVPSRPQSCINNNHRRHGYTSSLELPPNLLNPKKHPDLMEDQVPRMSRPLLYKGTNF	420
DB	361	PVPSRPQSCINNNHRRHGYTSSLELPPNLLNPKKHPDLMEDQVPRMSRPLLYKGTNF	420
QY	421	THLVADRYTGIDGATYVTLFTGTGQWLLKAVSGLPWWHLLEQLFQEPKRSIYLSQS	480
DB	421	THLVADRYTGIDGATYVTLFTGTGQWLLKAVSGLPWWHLLEQLFQEPKRSIYLSQS	480
QY	481	KLLFAGRSQSLVOLPVADCIKRRSCADCYLAADPYCANSVTSRCVAVGGHFSLLIQH	540
DB	481	KLLFAGRSQSLVOLPVADCIKRRSCADCYLAADPYCANSVTSRCVAVGGHFSLLIQH	540
QY	541	VMTSDTSGICNLKSGIKSGIPKNIIVAGTDLVLPCHLSNLA-----LPDSN	590
DB	541	VMTSDTSGICNLKSGIKSGIPKNIIVAGTDLVLPCHLSNLA-----LPDSN	590
QY	591	P 591	
DB	600	P 600	

RESULT 6  
 AAE18214  
 ID AAE18214 standard: Protein; 833 AA.  
 AC AAE18214;  
 XX  
 XX 07-MAY-2002 (first entry)  
 DE Human MO15b protein.  
 XX  
 XX Secreted molecule; MO15b protein; MOlx; cardiomyopathy; atherosclerosis;  
 KM diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;  
 KM liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;  
 KM cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;  
 KM immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;  
 KM HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;  
 KM haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;  
 KM Parkinson's disease; Huntington's disease; muscular disease; stress;  
 KM ocular disease; growth disorder; depression; epilepsy; contraceptive;  
 KM vulnery; osteopathic; haemostatic; tranquilliser; antidepressant;  
 KM analgesic; vasotropic; hypotensive; gene therapy; chromosome 2.  
 XX  
 XX Homo sapiens.

XX MO200206339-A2.  
 XX 24-JAN-2002.  
 XX 03-JUL-2001: 2001MO-US21249.  
 XX 03-JUL-2000: 2000US-215854P.  
 XX 03-JUL-2000: 2000US-215856P.  
 XX 03-JUL-2000: 2000US-215902P.  
 XX 07-JUL-2000: 2000US-216585P.  
 XX 07-JUL-2000: 2000US-216586P.  
 XX 07-JUL-2000: 2000US-216722P.  
 XX 17-JUL-2000: 2000US-218642P.  
 XX 17-JUL-2000: 2000US-218992P.  
 XX 27-JUL-2000: 2000US-221285P.  
 XX 14-FEB-2001: 2001US-268734P.  
 XX (CURA-) CURAGEN CORP.  
 XX Spaderna SK, Tchernev V, Liu X, Shenoy S, Spytek K, Zerhusen B;  
 XX Paturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;  
 XX Alsbrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;  
 XX Padigaru M;  
 XX WPI, 2002-155038/20.  
 XX N-PSDB: AAD28948.  
 XX Nucleic acids encoding secreted polypeptides, designated MOLX  
 XX polypeptides, useful for treating a MOLX-associated disorder, e.g.  
 XX cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -  
 XX claim 1: Page 61-62; 223pp; English.  
 XX The patent discloses nucleic acid sequences encoding novel secreted  
 XX molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL  
 XX protein where X is an integer from 1 to 8). Sequences of the invention  
 XX are useful for treating or preventing a MOLX-associated disorder in  
 XX humans. They are useful for treating or preventing cardiomyopathy,  
 XX atherosclerosis and disorders related to cell signal processing and  
 XX metabolic pathway modulation. The MOLX antibodies are useful for  
 XX treating or preventing diabetes and disorders related to cell signal  
 XX processing and metabolic pathway modulation. MOLX sequences are useful  
 XX for the treatment or diagnosis of other MOLX-associated disorders, e.g.  
 XX chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,  
 XX scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral  
 XX thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune  
 XX diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,  
 XX bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,  
 XX haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,  
 XX ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 XX ocular disease, muscular diseases, growth disorders, loss of libido,  
 XX stress, depression, pain and epilepsy. They are useful for preventing  
 XX chemotherapy side effects and as contraceptives. Sequences of the  
 XX invention are also useful for gene therapy. The present sequence  
 XX is human semaphorin 4C-like protein, MOL5b. MOL5b gene is localised  
 XX on chromosome 2.  
 XX SQ Sequence 833 AA;  
 XX  
 XX Query Match 94.9%; Score 3055.5; DB 23; Length 833;  
 XX . Best Local Similarity 94.7%; Pred. No. 1.7e-300;  
 XX Matches 569; Conservative 4; Mismatches 17; Indels 11; Gaps 2;  
 XX  
 XX 1 MAPHNAVLLAARLMGIGAEVWNLVPRKTVSSGELATVYRRFSQTGIQDFLTTLTE 60  
 XX 1 MAPHNAVLLAARLMGIGAEVWNLVPRKTVSSGELATVYRRFSQTGIQDFLTTLTE 60  
 XX 1 MAPHNAVLLAARLMGIGAEVWNLVPRKTVSSGELATVYRRFSQTGIQDFLTTLTE 60  
 XX 61 PTGLLYVGAREALFAFMSKEALEGALISWEAPVEKTECIQKKNQTECFNFIREFLQPY 120  
 XX 61 PTGLLYVGAREALFAFMSKEALEGALISWEAPVEKTECIQKKNQTECFNFIREFLQPY 120  
 XX 61 PTGLLYVGAREALFAFMSKEALEGALISWEAPVEKTECIQKKNQTECFNFIREFLQPY 120  
 XX 61 PTGLLYVGAREALFAFMSKEALEGALISWEAPVEKTECIQKKNQTECFNFIREFLQPY 120  
 XX 121 NASHLYVCGTYAFQPRCTYVNMILFTLEHGFEDGKCKCPYDPAKHAGLLVDGELYSAT 180  
 XX 121 NASHLYVCGTYAFQPRCTYVNMILFTLEHGFEDGKCKCPYDPAKHAGLLVDGELYSAT 180

DB 121 NASHLYVCGTYAFQPRCTYVNMILFTLEHGFEDGKCKCPYDPAKHAGLLVDGELYSAT 180  
 QY 181 LNNEFLGTEPTILNNMGPBHSMTKTEYLAFWLNHEHFGSAVVPSPVSFTGDDDKKYFFER 240  
 DB 181 LNNEFLGTEPTILNNMGPBHSMTKTEYLAFWLNHEHFGSAVVPSPVSFTGDDDKKYFFER 240  
 QY 241 ERAVESDCYAEQVAVARVARYCKGDMGARTIQRKWTTFKLARLACAPNMQLYFNLOAM 300  
 DB 241 ERAVESDCYAEQVAVARVARYCKGDMGARTIQRKWTTFKLARLACAPNMQLYFNLOAM 300  
 QY 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEBEIQRVFESPYEYHEBAQKWRD 360  
 DB 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEBEIQRVFESPYEYHEBAQKWRD 360  
 QY 361 PVSPRPGSCINNMHRHGGTSSLELPDNLNFKKHPIMEBOVGRWSPRLLYKKGCTNF 420  
 DB 361 PVSPRPGSCINNMHRHGGTSSLELPDNLNFKKHPIMEBOVGRWSPRLLYKKGCTNF 420  
 QY 421 THLYADRVTLGDATYTVLFTIGDGMVLKAVSLGPMVHLIEBLQFPDEPMRSLVLSOS 480  
 DB 421 THLYADRVTLGDATYTVLFTIGDGMVLKAVSLGPMVHLIEBLQFPDEPMRSLVLSOS 480  
 QY 481 KKLFPAGSSQVQLPVADCIKTRSCADCVLANDPYCAMSVNTSRCAVAGHFGSLILIOH 540  
 DB 481 KKLFPAGSSQVQLPVADCIKTRSCADCVLANDPYCAMSVNTSRCAVAGHFGSLILIOH 540  
 QY 541 VMTSDRSGICNLNMGSKYRSCADCVLANDPYCAMSVNTSRCAVAGHFGSLILIOH 590  
 DB 541 VMTSDRSGICNLNMGSKYRSCADCVLANDPYCAMSVNTSRCAVAGHFGSLILIOH 590  
 QY 591 P 591  
 DB 600 P 600  
 DB 600 P 600  
 RESULT 7  
 AAE18215  
 ID AAE18215 standard; Protein; 833 AA.  
 NC AAE18215;  
 XX 07-MAY-2002 (first entry)  
 DT  
 XX Human MOL5c protein.  
 XX  
 XX Secreted molecule; MOL5c protein; MOLX; cardiomyopathy; atherosclerosis;  
 KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;  
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;  
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;  
 KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;  
 KW HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;  
 KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;  
 KW ocular disease; growth disorder; depression; epilepsy; antidepressant;  
 KW vulnary; osteopathic; haemostatic; tranquiliser; antiplatelet;  
 KW analgesic; vasotropic; hypotensive; gene therapy; chromosome 2.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..20  
 FT /label= signal\_peptide  
 FT Protein 21..833  
 FT /label= Mature\_MOL5c\_protein  
 XX  
 XX MO200206339-A2.  
 XX  
 XX 24-JAN-2002.  
 XX  
 XX 03-JUL-2001: 2001MO-US21249.  
 XX  
 XX 03-JUL-2000: 2000US-215854P.



PR 03-JUL-2000; 2000US-215856P.  
 PR 03-JUL-2000; 2000US-215902P.  
 PR 07-JUL-2000; 2000US-216585P.  
 PR 07-JUL-2000; 2000US-216586P.  
 PR 07-JUL-2000; 2000US-216722P.  
 PR 17-JUL-2000; 2000US-218623P.  
 PR 17-JUL-2000; 2000US-218992P.  
 PR 27-JUL-2000; 2000US-221285P.  
 PR 14-FEB-2001; 2001US-268734P.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Spaderna SK, Tchernev V, Liu X, Shenoy S, Spytek K, Zernhusen B;  
 PI Paturajan M, Taupier RJ, Rastell L, Grosse WM, Szelekes ES;  
 PI Alsobrook J, Iepley DM, Shen L, Burgess CE, Shinkets RA;  
 PI Padigaru M;  
 XX  
 DR WPI: 2002-155038/20.  
 DR N-PSDB: AAD28949.  
 XX  
 PT Nucleic acids encoding secreted polypeptides, designated MOIX  
 PT polypeptides, useful for treating a MOIX-associated disorder, e.g.  
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -  
 XX  
 PS Claim 1; Page 64; 223pp; English.  
 XX  
 CC The patent discloses nucleic acid sequences encoding novel secreted  
 CC molecule (MOI) polypeptides, designated MOIX polypeptides (i.e. a MOI  
 CC protein where X is an integer from 1 to 8). Sequences of the invention  
 CC are useful for treating or preventing a MOIX-associated disorder in  
 CC humans. They are useful for treating or preventing cardiomyopathy,  
 CC atherosclerosis and disorders related to cell signal processing and  
 CC metabolic pathway modulation. The MOIX antibodies are useful for  
 CC treating or preventing diabetes and disorders related to cell signal  
 CC processing and metabolic pathway modulation. MOIX sequences are useful  
 CC for the treatment or diagnosis of other MOIX-associated disorders, e.g.  
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,  
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral  
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune  
 CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,  
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,  
 CC haemolytic, thrombolytic, haemostatic diseases, thrombocytopaenia,  
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC ocular disease, muscular diseases, growth disorders, loss of libido,  
 CC stress, depression, pain and epilepsy. They are useful for preventing  
 CC chemotherapy side effects and as contraceptives. Sequences of the  
 CC invention are also useful for gene therapy. The present sequence  
 CC is human semaphorin 4C-like protein, MOI5C. MOI5C gene is localised  
 CC on chromosome 2.  
 CC  
 XX  
 SQ Sequence 833 AA;  
 Query Match 94.9%; Score 3055.5; DB 23; Length 833;  
 Best Local Similarity 94.7%; Pred. No. 1.7e-300;  
 Matches 569; Conservative 4; Mismatches 17; Indels 11; Gaps 2;

DB 241 ERAVESDCTAEQVAVARAVAVCKGDMGARTLQRTKMTFLKARLACSPANMOLYFNOLQAM 300  
 QY HTLADTSWNTTFFGVFOAQMDYLSAICETOLEEIORVEGPYKEHEEAKMDRYTD 360  
 DB 301 HTLADTSWNTTFFGVFOAQMDYLSAICETOLEEIORVEGPYKEHEEAKMDRYTD 360  
 DB 301 HTLADTSWNTTFFGVFOAQMDYLSAICETOLEEIORVEGPYKEHEEAKMDRYTD 360  
 QY 361 PVSPRPSCINNNHRRHGYTSSLEPDLNLFYKHPLEBOYQPRMSRPLVYKGTNF 420  
 DB 361 PVSPRPSCINNNHRRHGYTSSLEPDLNLFYKHPLEBOYQPRMSRPLVYKGTNF 420  
 QY 421 THLVADRYTGIDGATTYVLFITGIDGRLKAVSICPWHVLIIEQLDEPMSRLVYSQS 480  
 DB 421 THLVADRYTGIDGATTYVLFITGIDGRLKAVSICPWHVLIIEQLDEPMSRLVYSQS 480  
 QY 481 KKLIFAGRSQVLQVPVADCKKRYSCADCVLARPCTCAMSVMYTRCAVAGSGSLIQH 540  
 DB 481 KKLIFAGRSQVLQVPVADCKKRYSCADCVLARPCTCAMSVMYTRCAVAGSGSLIQH 540  
 QY 541 VMTSDTSGICNLKRSKIOSGFPKNTVYVACTDLVLPCHLSSNLA-----LPDSN 590  
 DB 541 VMTSDTSGICNLKRSK-KVRPTPKNTVYVACTDLVLPCHLSSNLAHARWTRGDRDLPAEQ 599  
 QY 591 P 591  
 DB 600 P 600  
 RESULT 8  
 ABB98402  
 ID ABB98402 standard; Protein: 805 AA.  
 XX  
 XX ABB98402;  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Human NOV2, a semaphorin like protein.  
 XX  
 KW Human; NOV2; cytosolic; Cardiac; Antiinflammatory; Immunosuppressive;  
 KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;  
 KW Antithrombotic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;  
 KW Antibacterial; Viucide; Antiparasitic; Relaxant; Anticonvulsant;  
 KW Gene therapy; NOV; cancer; heart disease; Inflammation;  
 KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes;  
 KW obesity; asthma; Iga nephropathy; cirrhosis; arthritis;  
 KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;  
 KW wasting disorder; semaphorin-like protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20025704-A2.  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 09-JAN-2002; 2002WO-US00554.  
 XX  
 PR 09-JAN-2001; 2001US-260417P.  
 PR 10-JAN-2001; 2001US-260831P.  
 PR 28-FEB-2001; 2001US-272386P.  
 PR 09-MAR-2001; 2001US-274876P.  
 PR 18-APR-2001; 2001US-284704P.  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Padigaru M, Li L, Zernhusen BD, Casman SJ, Shenoy S, Spytek KA;  
 PI Zhong M, Gangoli EA, Burgess CE, Paturajan M, Vernet CAM;  
 PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;  
 PI Alsobrook J, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;  
 PI MacDougal J, Malysankar U, Millet I, Peyman J, Smithson G;  
 PI Gunther E, Stone DJ;  
 XX  
 DR WPI: 2002-590674/63.  
 DR N-PSDB: ABB85379.

XX NOVX polypeptides and encoding polynucleotides, useful for preventing  
 PT or treating NOVX-associated disorders e.g. cancer, inflammation, or  
 PT Alzheimer's disease, and in chromosome mapping, tissue typing or  
 PT pharmacogenomics -  
 XX  
 PS Claim 1, Page 20; 358pp; English.  
 CC The present sequence is the protein sequence for a NOV protein. The  
 CC NOV proteins and coding sequences are useful for treating or preventing  
 CC NOV-associated disorders or in the manufacture of a medicament for  
 CC treating the disorders, such as cancer, heart disease, inflammation,  
 CC autoimmune disorders, allergies, blood disorders, AIDS, diabetes,  
 CC obesity, asthma, Iga nephropathy, cirrhosis, arthritis, Alzheimer's  
 CC disease, infections (e.g. bacterial, viral, parasitic), stroke, muscular  
 CC dystrophy, epilepsy, and other wasting disorders associated with chronic  
 CC diseases. NOV2 is a semaphorin like protein.  
 XX  
 SQ Sequence 805 AA;  
 Query Match 90.9%; Score 2927.5; DB 23; Length 805;  
 Best Local Similarity 91.7%; Pred. No. 1.6e-287;  
 Matches 551; Conservative 2; Mismatches 9; Indels 39; Gaps 3;  
 QY 1 MAPRNAWLLAARLKGIGAEVWNNLYPRKTVSSGELATVYRRSQTGIODFLTITLRE 60  
 DB 1 MAPRNAWLLAARLKGIGAEVWNNLYPRKTVSSGELATVYRRSQTGIODFLTITLRE 60  
 QY 61 PTGLTYVGAREALFAFMSMALELOGAISWEAPVEKTECIQKKNQTCFNFIRLOPY 120  
 DB 61 PTGLTYVGAREALFAFMSMALELOGAISWEAPVEKTECIQKKNQTCFNFIRLOPY 120  
 QY 121 NASHLYVCGTYAFOPKCTVYNNMLTFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180  
 DB 121 NASHLYVCGTYAFOPKCTVYNNMLTFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180  
 QY 181 LNNFLCTEPTIILRNKGPBHSMTETYLAFWLNPHFPGSAVYVESVGSFTGDDDKYFFER 240  
 DB 181 LNNFLCTEPTIILRNKGPBHSMTETYLAFWLNPHFPGSAVYVESVGSFTGDDDKYFFER 240  
 QY 241 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 300  
 DB 241 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 300  
 QY 213 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 272  
 DB 213 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 272  
 QY 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 360  
 DB 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 360  
 QY 273 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 332  
 DB 273 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 332  
 QY 361 PVPSPRPGSCINNMHRHGGYTSLELPOILNFVKKHPIMEBOVGPRMSRPILLYKKGTFN 420  
 DB 361 PVPSPRPGSCINNMHRHGGYTSLELPOILNFVKKHPIMEBOVGPRMSRPILLYKKGTFN 420  
 QY 333 PVPSPRPGSCINNMHRHGGYTSLELPOILNFVKKHPIMEBOVGPRMSRPILLYKKGTFN 392  
 DB 333 PVPSPRPGSCINNMHRHGGYTSLELPOILNFVKKHPIMEBOVGPRMSRPILLYKKGTFN 392  
 QY 421 THLVADRYTGLDGAATYTVLFTIGTGDMWLLKAVSLGPVHLLIELOLFDQEPKSLVLSOS 480  
 DB 421 THLVADRYTGLDGAATYTVLFTIGTGDMWLLKAVSLGPVHLLIELOLFDQEPKSLVLSOS 480  
 QY 393 THLVADRYTGLDGAATYTVLFTIGTGDMWLLKAVSLGPVHLLIELOLFDQEPKSLVLSOS 452  
 DB 393 THLVADRYTGLDGAATYTVLFTIGTGDMWLLKAVSLGPVHLLIELOLFDQEPKSLVLSOS 452  
 QY 481 KLLILAGRSQVQLPVADCIYRSCADCVLARDPYCAMSVTSRCAVAGHPSLLIHH 540  
 DB 481 KLLILAGRSQVQLPVADCIYRSCADCVLARDPYCAMSVTSRCAVAGHPSLLIHH 540  
 QY 453 KLLILAGRSQVQLPVADCIYRSCADCVLARDPYCAMSVTSRCAVAGHPSLLIHH 512  
 DB 453 KLLILAGRSQVQLPVADCIYRSCADCVLARDPYCAMSVTSRCAVAGHPSLLIHH 512  
 QY 541 VMTSDTSGICNLGKSGIKSGPKNTVYAGTDVLPCHLSSNLA-----LPDSN 590  
 DB 541 VMTSDTSGICNLGKSGIKSGPKNTVYAGTDVLPCHLSSNLA-----LPDSN 590  
 QY 591 P 591  
 DB 591 P 591  
 QY 572 P 572  
 DB 572 P 572  
 RESULT 9  
 ABB97963  
 ID ABB97963 standard; Protein; 886 AA.  
 XX

AC ABB97963;  
 XX  
 XX 06-SEP-2002 (first entry)  
 DT  
 XX  
 DE Human protein sequence #30.  
 XX  
 XX Human; brain; tonsil; hippocampus; foetal brain; diagnosis.  
 KW  
 XX Homo sapiens.  
 OS  
 XX MO200252005-A1.  
 PN  
 XX 04-JUL-2002.  
 PD  
 XX 20-DEC-2001; 2001WO-JP11217.  
 PF  
 XX 22-DEC-2000; 2000JP-0389742.  
 PR  
 XX (KAZU-) KAZUSA DNA RES INST FOUND.  
 PA (CELE-) CELESTAR LEXICO-SCI LTD.  
 XX  
 XX Ohara O, Nagase T, Nakajima D;  
 PI WPI, 2002-500762/53.  
 DR N-PSDB; ABB93983.  
 XX  
 XX Genes and their expression products cloned from human cDNA libraries  
 PT for treatment and diagnosis of diseases associated with their  
 PT expression -  
 XX  
 XX Claim 1(a); Page 119-125; 238pp; Japanese.  
 PS  
 XX The invention relates to DNA encoding polypeptides directly cloned from  
 CC cDNA libraries originating in adult whole brain, human tonsil, human  
 CC adult hippocampus and human foetal whole brain. Polypeptides and  
 CC polynucleotides of the invention may be used in the investigation of  
 CC differential expression of the DNA sequences in normal subjects and  
 CC disease patients. They may also be used in the production of antibodies,  
 CC oligonucleotide probes and DNA chips for diagnosis and identification  
 CC of drugs for treatment of diseases with which the DNA sequences are  
 CC associated. The sequences given in records ABB97934-ABB97964 represent  
 CC human proteins of the invention.  
 CC  
 SQ Sequence 886 AA;  
 Query Match 90.6%; Score 2918; DB 23; Length 886;  
 Best Local Similarity 92.2%; Pred. No. 1.7e-286;  
 Matches 554; Conservative 5; Mismatches 22; Indels 20; Gaps 5;  
 QY 8 W-LIARLKGIGAEVWNNLYPRKTVSS-----GELATVYRRSQTGIODFLTITLRE 60  
 DB 8 W-LIARLKGIGAEVWNNLYPRKTVSS-----GELATVYRRSQTGIODFLTITLRE 60  
 QY 56 WHLVAV--WGAGSRGARLRAVEFGSCPSAAMLTPALATVYRRSQTGIODFLTITLRE 113  
 DB 56 WHLVAV--WGAGSRGARLRAVEFGSCPSAAMLTPALATVYRRSQTGIODFLTITLRE 113  
 QY 61 PTGLTYVGAREALFAFMSMALELOGAISWEAPVEKTECIQKKNQTCFNFIRLOPY 120  
 DB 61 PTGLTYVGAREALFAFMSMALELOGAISWEAPVEKTECIQKKNQTCFNFIRLOPY 120  
 QY 114 PTGLTYVGAREALFAFMSMALELOGAISWEAPVEKTECIQKKNQTCFNFIRLOPY 173  
 DB 114 PTGLTYVGAREALFAFMSMALELOGAISWEAPVEKTECIQKKNQTCFNFIRLOPY 173  
 QY 121 NASHLYVCGTYAFOPKCTVYNNMLTFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180  
 DB 121 NASHLYVCGTYAFOPKCTVYNNMLTFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 180  
 QY 174 NASHLYVCGTYAFOPKCTVYNNMLTFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 233  
 DB 174 NASHLYVCGTYAFOPKCTVYNNMLTFTLEHGEFEDGKGCPRYPANAGHLLVDGELYSAT 233  
 QY 181 LNNFLCTEPTIILRNKGPBHSMTETYLAFWLNPHFPGSAVYVESVGSFTGDDDKYFFER 240  
 DB 181 LNNFLCTEPTIILRNKGPBHSMTETYLAFWLNPHFPGSAVYVESVGSFTGDDDKYFFER 240  
 QY 234 LNNFLCTEPTIILRNKGPBHSMTETYLAFWLNPHFPGSAVYVESVGSFTGDDDKYFFER 293  
 DB 234 LNNFLCTEPTIILRNKGPBHSMTETYLAFWLNPHFPGSAVYVESVGSFTGDDDKYFFER 293  
 QY 241 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 300  
 DB 241 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 300  
 QY 294 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 353  
 DB 294 ERAVESDCYAEQVAVARAVCKDGMGARTLQKMTTFPKARLACAPMMOLYFNLOLM 353  
 QY 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 360  
 DB 301 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 360  
 QY 354 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 413  
 DB 354 HTLQDTSMHNTTFFGVFOAQMGDMYLSAICEYOLEIEIQVFEFGPYKEHYEEAOKMDRYTD 413

QY 361 PVSPRPGSCINNNHRRHGTSSLELDPNIIINFKKHPLEMEQVGPMSRPLVKKGTNF 420  
DB 414 PVSPRPGSCINNNHRRHGTSSLELDPNIIINFKKHPLEMEQVGPMSRPLVKKGTNF 473  
QY 421 THLVADRVTVGLDGAATYVLTFTGTGDMILKAVSLGPPWHLLIEQLPQEPMBRLVLSQS 480  
DB 474 THLVADRVTVGLDGAATYVLTFTGTGDMILKAVSLGPPWHLLIEQLPQEPMBRLVLSQS 533  
QY 481 KKLIFAGRSQSLVOLPVADCIKYSKADCVIARDPYCAMSVNTSRCAVAGSHSGLLIQH 540  
DB 534 KKLIFAGRSQSLVOLPVADCIKYSKADCVIARDPYCAMSVNTSRCAVAGSHSGLLIQH 593  
QY 541 VMTSDTSGICNLKRSKIQSGPXPKNITVYAGTDLVLPCHLSSNLA-----LPDSN 590  
DB 594 VMTSDTSGICNLKRSK--KVRPTPKNITVYAGTDLVLPCHLSSNLAHAWTFGGRDLPAEQ 652  
QY 591 P 591  
DB 653 P 653

RESULT 10  
AAB61238  
ID AAB61238 standard; Protein: 834 AA.  
XX  
AC AAB61238;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE Murine M-Sema-F protein.  
XX  
KW Mouse; M-Sema-F; INTERCEPT 217; INTERCEPT 297; TANGO 276;  
KW TANGO 292; TANGO 325; TANGO 331; TANGO 332; cycostatic; antiinflammatory;  
KW antiarrhythmic; antipsoritic; gene therapy; cancer;  
KW inflammatory disorder; cardiac disorder; arrhythmia; skin disorder;  
KW psoriasis.  
XX  
OS Mus sp.  
XX  
PN WO200100638-A2.  
XX  
PD 04-JAN-2001.  
XX  
PF 16-JUN-2000; 2000MO-US16658.  
XX  
PR 29-JUN-1999; 99US-0342364.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Kitst SJ, Holtzman DA, Fraser CC, Sharp JD, Barnes TS;  
XX  
DR WPI: 2001-061966/07.  
XX  
N-PSDB; AAF29461.  
XX  
PT Isolated human proteins are used for diagnosis, treatment and  
PT prevention of cancers, inflammatory disorders, cardiac disorders e.g.  
PT arrhythmia, and skin disorders e.g. psoriasis -  
XX  
XX  
PS Disclosure; Fig 3F-3H; 372pp; English.  
XX  
CC The present sequence is given in a specification relating to isolated  
CC human proteins designated INTERCEPT 217, INTERCEPT 297, TANGO 276,  
CC TANGO 292, TANGO 325, TANGO 331 and TANGO 332. These proteins are useful  
CC as modulating agents or as targeting agents for developing agents to  
CC regulate cellular processes e.g. growth, proliferation, survival,  
CC differentiation and activity of human tissues. Diseases which can be  
CC diagnosed, prevented and treated by administration of these polypeptides,  
CC their nucleic acids and modulators include cancers, inflammatory  
CC disorders, cardiac disorders e.g. arrhythmia, and skin disorders e.g.  
CC psoriasis. Nucleic acids encoding the isolated proteins can be used to  
CC express the proteins in a host cell in gene therapy applications.  
CC Antisense molecules or ribozymes can be used to inhibit expression  
CC of the proteins in target cells. Fragments of the nucleic acid molecules

CC encoding the isolated proteins are used as hybridisation probes and as  
CC polymerase chain reaction (PCR) primers.  
XX  
SQ Sequence 834 AA;  
QY Query Match 84.5%; Score 2720; DB 22; Length 834;  
DB Best Local Similarity 83.1%; Pred. No. 2e-266;  
DB Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;  
QY 1 MAPHAWVLLAARLWGLGIGAEVMMNLYPRRTVSSGELATVRRPSQIGDFFLTITTE 60  
DB 1 MAPHAWVLLAAGLWGLGIGAEVMMNLYPRRTVSSGELATVRRPSQIGDFFLTITTE 60  
QY 61 PTGLIYGARBALFAFSMEALELQGAISMEAPVEKTEICIOGKNNOTECNFIFLQPY 120  
DB 61 HSGLIYGARBALFAFSMEALELQGAISMEAPVEKTEICIOGKNSNTECFNIFLQPY 120  
QY 121 NASHLYVCGTYAFQPKCTYVNMVLTFTLEHGEFEDGKGCYPDPAGHAGLLVDGELYAT 180  
DB 121 NSSHLVYCGTYAFQPKCTYVNMVLTFTLEHGEFEDGKGCYPDPAGHAGLLVDGELYAT 180  
QY 181 LNNFLGTEPITLLRNNGPHHSMKTEYLAFMLNPHFVGSAYVPSVGSFTGDDKTYEPPR 240  
DB 181 LNNFLGTEPITLLRNNGPHHSMKTEYLAFMLNPHFVGSAYVPSVGSFTGDDKTYEPPR 240  
QY 241 ERAVESDCYAEQVAVARAVCKGDMGARTLQKRTTFELKARLACSAFNMQLYFNQLAM 300  
DB 241 ERAVEYDCYSQVAVARAVCKGDMGARTLQKRTTFELKARLACSAFNMQLYFNQLAM 300  
QY 301 HTLQDTSMHNTTFEGVFOAQWQMDTLAICEYQLEIQRVVEGPKYKEYHEEAQKMDRYTD 360  
DB 301 HTLRGASWHNTTFEGVFOARWQMDTLAICEYQLEIQRVVEGPKYKEYHEEAQKMDRYTD 360  
QY 361 PVSPRPGSCINNNHRRHGTSSLELDPNIIINFKKHPLEMEQVGPMSRPLVKKGTNF 420  
DB 361 PVSPRPGSCINNNHRRHGTSSLELDPNIIINFKKHPLEMEQVGPMSRPLVKKGTNF 420  
QY 421 THLVADRVTVGLDGAATYVLTFTGTGDMILKAVSLGPPWHLLIEQLPQEPMBRLVLSQS 480  
DB 421 THLVADRVTVGLDGAATYVLTFTGTGDMILKAVSLGPPWHLLIEQLPQEPMBRLVLSQS 480  
QY 481 KKLIFAGRSQSLVOLPVADCIKYSKADCVIARDPYCAMSVNTSRCAVAGSHSGLLIQH 539  
DB 481 KKLIFAGRSQSLVOLPVADCIKYSKADCVIARDPYCAMSVNTSRCAVAGSHSGLLIQH 540  
QY 540 VMTSDTSGICNLKRSKIQSGPXPKNITVYAGTDLVLPCHLSSNLA-----LPD 588  
DB 540 VMTSDTSGICNLKRSKIQSGPXPKNITVYAGTDLVLPCHLSSNLAHAWTFGGRDLPA 598  
QY 589 SNP 591  
DB 599 BGP 601

RESULT 11  
ABU60951  
ID ABU60951 standard; Protein: 785 AA.  
XX  
XX  
AC ABU60951;  
XX  
DT 08-MAY-2003 (first entry)  
XX  
DE Lung specific protein (LSP) #54.  
XX  
KW Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;  
KW cancer monitoring; cancer staging; cancer imaging; lung cancer;  
KW non-cancerous diseases of the lung; transgenic animal.  
XX  
OS Homo sapiens.  
XX  
PN WO200268633-A2.  
XX  
PD 06-SEP-2002.

XX 21-NOV-2001; 2001MO-US43612.  
PE  
XX  
PR 22-NOV-2000; 2000US-25250P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI MacIne RA, Reclpon H, Chen S, Sun Y, Liu C;  
DR WPI; 2002-713376/77.  
XX  
PT New isolated human nucleic acid molecule and polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating lung  
PT cancer and non-cancerous diseases of the lung .  
XX  
PS Claim 11; Page 345-348; 389pp; English.  
XX

The invention describes an isolated human nucleic acid (I) encoding any of 120 10-1533 residue amino acid sequences (S1), given in the specification, comprising any of 164 179-12421 base pair sequences (S2), given in the specification. The methods and compositions of the present invention are useful for identifying, diagnosing, monitoring, staging, imaging and treating lung cancer and non-cancerous diseases of the lung. They are also used for identifying lung tissue, monitoring and identifying and/or designing antagonists of the polypeptide of the invention, gene therapy, production of transgenic animals and production of engineered lung tissue for treatment and research. This is the amino acid sequence of a lung specific nucleic acid.

Sequence 785 AA:

Query Match	91.9%	Score 2635	DB 23	Length 785
Best Local Similarity	80.0%	Pred. No. 7, 8e-258		
Matches	495	Conservative	2	Mismatches 7
				Indels 46
				Gaps 4

  

QY	67	ISWEAPVEKTECIQKGNKNN-----QTECPN	112
DB	4	ISWEAPVEKTECIQKGNKNNQWATLLLVLPDVSBSGSRALNTNTPRECPQECN	63
QY	113	FIRFQAPRNASHLYVCGTAAFOPKCTYNNMLFETLEHGEFEGEGKGCPRDPKGNHGLLV	172
DB	64	FIRFQAPRNASHLYVCGTAAFOPKCTYNNMLFETLEHGEFEGEGKGCPRDPKGNHGLLV	123
QY	173	DGEIYSATLNNPLGTEPIILRNMGPHNSKTEYLFWILNEPHFVSAAVYVESVSGFTGDD	232
DB	124	DGEIYSATLNNPLGTEPIILRNMGPHNSKTEYLFWILNEPHFVSAAVYVESVSGFTGDD	183
QY	233	DKYVFFERRAVESCYAEQVAVARARCKGMSGARTLQRIKNTTFPKARIKIASAPNMQL	292
DB	184	DKYVFFERRAVESCYAEQVAVARARCKGMSGARTLQRIKNTTFPKARIKIASAPNMQL	243
QY	293	YFNQLQAMHTLQDTSMNHTTFEGVFOAOWMDYTSACEYOLEEIQREFEGPKYHEBA	352
DB	244	YFNQLQAMHTLQDTSMNHTTFEGVFOAOWMDYTSACEYOLEEIQREFEGPKYHEBA	303
QY	353	QKMDRYTPRVSPRGSGCINNNHNRHNGVTSLELDNLIAPFKKIRYMEQVGRPSRDL	412
DB	304	QKMDRYTPRVSPRGSGCINNNHNRHNGVTSLELDNLIAPFKKIRYMEQVGRPSRDL	363
QY	413	LVKKGTFNTHLVADRVTLGDGATYVLEIGTGDGMLKAVSLGAPWNLITEEQLFQDEBM	472
DB	364	LVKKGTFNTHLVADRVTLGDGATYVLEIGTGDGMLKAVSLGAPWNLITEEQLFQDEBM	423
QY	473	RSLYLSQSK-KLTPAGRSQVQLQVPLVADCIKTRSCADCVLADRPICAMSVNTSRCAVGG	531
DB	424	RSLYLSQSKVLLTPAGRSQVQLQVPLVADCIKTRSCADCVLADRPICAMSVNTSRCAVGG	483
QY	532	HFGSLILQHVMTSDPTSGICINLRSKGIQSGXPKNITVYAGTDLVLRCHLSSNLA-----	585
DB	484	HSGSLILQHVMTSDPTSGICINLRSK-KVRPFPKNTTVVAGTDLVLRCHLSSNLAIRKTF	542
QY	586	-----LPPDSNP 591	

Db 543 GGRDLPANQP 552

RESULT 12  
ABJ18435  
ID ABJ18435 standard; Protein; 785 AA.  
XX ABJ18435;  
AC  
XX  
XX  
DT 30-JAN-2003 (first entry)  
XX  
DE Breast specific related amino acid sequence SEQ ID No 244.  
XX  
XX Cytostatic; BSP-agonist; BSP-antagonist; vaccine; gene therapy; cancer;  
KW metastatic; breast cancer; breast specific; human.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO200277232-A2.  
PN  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 21-NOV-2001; 2001WO-US43815.  
PF  
XX  
XX 22-NOV-2000; 2000US-252509P.  
PR  
XX  
XX (DIAD-) DIADEXUS INC.  
PA  
XX  
XX Salceda S, Macina RA, Recipon H, Pluta J, Sun Y, Liu C;  
XX

```

XX New isolated nucleic acid molecule, useful for treating breast cancer,
PT and diagnosing or monitoring the presence of metastases of breast
PT cancer in a patient
PS
PS
XX Claim 11; Page 358-361; 377pp; English.
XX
CC The invention relates to a novel isolated nucleic acid molecule
CC comprising: a sequence encoding a sequence comprising 11-1518 amino
CC acids; a sequence comprising 190-8144 bp; or a sequence that selectively
CC hybridises to, or having at least 60% identity with the 11-1518 amino
CC acid or 190-8144 nt sequence. The polypeptide and the nucleic acid are
CC useful for treating breast cancer, and diagnosing or monitoring the
CC presence of metastases of breast cancer in a patient. The polynucleotides
CC of the invention can be used to treat disorders by gene therapy. This
CC sequence represents a breast specific related polypeptide of the
CC invention.
XX
XX Sequence 785 AA;
SQ
Query Match 81.9%; Score 2635; DB 24; Length 785;
Best Local Similarity 90.0%; Pred. No. 7.8e-258;
Matches 495; Conservative 2; Mismatches 7; Indels 46; Gaps 4;
QY 87 ISMAPPEKTECLOKRNK-----QTECFN 112
DB 4 ISWAPPEKTECLOKRNNOVGAMTLVLPSQDVSSHSGRALNRRPFCOTECFN 63
QY 113 FIRELOPYNASHLYVCGTAFOPACTYVNNMLTFLEHGEEDGKCKPYPAKHAGLV 172
DB 64 FIRELOPYNASHLYVCGTAFOPRCTYVNNMLTFLEHGEEDGKCKPYPAKHAGLV 123
QY 173 DGEIYSATLNNFGLTEPIILRNNGPHHSKMTLEYLAFNLNPHFVGSAYVESVSGFTGD 232
DB 124 DGEIYSATLNNFGLTEPIILRNNGPHHSKMTLEYLAFNLNPHFVGSAYVESVSGFTGD 183
QY 233 DKYVEFPERARVESDCAEADVAVARAVVAVCGDNGGATLQIRKMTTEFKALACAPRWOL 292
DB 184 DKYVEFPERARVESDCAEADVAVARAVVAVCGDNGGATLQIRKMTTEFKALACAPRWOL 243
QY 293 YFNOLQAMHTLQDTSMTNNTFFVGFQAOQMDNYLSACEVQLSEIQVGFVGPKEVHEEA 352

```



DR N-PSDB: AAD08283.

Isolated nucleic acid molecule encoding a human secreted protein is used in preventing, treating or ameliorating a medical condition

Claim 11: Page 476-479; 553pp; English.

CC AAD08283-AAD08355 represent cDNAs corresponding to 23 human secreted  
CC protein genes, and AAE03818-AAE03870 represent the proteins they encode.  
CC AAE03871-AAE03896 represent human secreted protein fragments or variants.  
CC The secreted proteins and their genes are useful for preventing,  
CC treating or ameliorating medical conditions, e.g., by protein or gene  
CC therapy. Pathological conditions can be diagnosed by determining the  
CC amount of the new protein in a sample or by determining the presence of  
CC mutations in the new genes. Specific uses are described for each of the  
CC 23 genes, based on the tissues in which they are most highly expressed,  
CC and include developing products for the diagnosis or treatment of  
CC proliferative disorders, cancer, tumours, foetal and developmental  
CC abnormalities, hematopoietic disorders, diseases of the immune system,  
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
CC allergies, neurological disorders (e.g., Alzheimer's disease,  
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
CC cardiovascular disorders, angiogenic disorders, kidney disorders,  
CC gastrointestinal disorders, pregnancy-related disorders, endocrine  
CC disorders, and infections. The proteins can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues, to identify their  
CC cognate ligands or binding partners, and in chemotaxis, and can be used  
CC as a food additive or preservative to modify storage properties.  
CC Antibodies specific for a protein of the invention can be used in  
CC alleviating symptoms associated with the disorders mentioned above, and  
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
CC immunosorbent assay (ELISA). The present sequence represents a human  
CC secreted protein of the invention.

SO Sequence 832 AA:

Query Match 38.5%; Score 1240; DB 22; Length 832;  
Best Local Similarity 45.4%; Pred. No. 4,3e-116;  
Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

QY 25 WLVLRKTYSSGLATVVRFRFQSGTIGDPLTITLTPGLLVGAREALFAFSMELAQ 84  
DB 37 WALSPRISLPLGSERRPLRPAEHSNTALLLSRDGRTLVGAREALFAFSMELAQ 96  
QY 65 GA-----ISWEAFVEKTECTIOGKNNQTECFNIFRLOPYNASHLYVCGTVAFOPCQTV 140  
DB 97 GGEYBELMGADAEKKQCSFKGKDPQRCQVYITLPLSGSHLFTGCTAAFSFPCITY 156  
QY 141 NMLTFTLEHG-----FEDGKCKCPYDPAKGHAGLLVDGLYSATLNNFLAGTEPTILRM 195  
DB 157 NNEFTLARDEKGNVLEDKGRCPCFDPNFKSTALVYDGLYGTGVSSFGQNDPAISRQ 216  
QY 196 GPHSHMKTEYLAFNLNRPVGSAYVPSVSGFTGDDKVFYFFERAVESCIYEQVYA 255  
DB 217 SLRPT-KTESSLNMLQDPAFVASAVTPSSLSGLQDDDKIYFFETGGEFFEFETVY 275  
QY 256 RVARVCCKDMDGARTLQKRTFTFKARLACSAFNNOLVFNOAHMTLQDT--SMHNTF 313  
DB 276 RIAKCKDEDEGERLQGRWTSFLNQLCSRFDDGFPFNVDYFTLSPSQDMRDLTF 335  
QY 314 FGVFOAQM--GDMYLSAICEYQLEIQRVVEGPEYKEHEADKMDRYDTPVSPRGSCI 371  
DB 336 YGVFTSQMHRGTTESSAVCVFTMKDQVRFSLYKEVNETQOMTVTHPVPTRPGACI 395  
QY 372 NMHHRHIGYTSLEPLDNLINVFYKHPMLMEQVGRKRSPLLYKKGTFNTHLVADRYGL 431  
DB 396 TNSAREKRIINSSLOLPDRVLFNLKDHFLMDQV--SNMLLLOQARQORAVARVPL 452  
QY 432 DGATVTVFVIGTGDMLKAVSLGPMVHLIELQLEFDD-EPMRSLVLSOSKLLFAGSRS 490  
DB 452

DB 453 H-HTVDVFLGCTGDGLHRAVSVGPRVHIIIEQLIFSSGQPVQNLDTHRGLLYAASHS 511  
QY 491 QLVQLPVDACIKYRSCADCVLADPVCAMSVNTRCAVAGHPGSLT---IQHWMTSDP 546  
DB 512 GVVQVPMANCSLRSRGDCDLARDPYCAMS--GSSCKHVSIVPOLATRWIODIGASA 569  
QY 547 SGICN-----LRGSKIQSGPP-KNITVVAAGTDLVLPCHLSSNIA 585  
DB 570 KDLCSASVSPSPFVPTGKPCQOVOPORNTVNTLACPLLSNIA 613

# RESULT 15

ID ABG64522 standard; Protein; 832 AA.  
AC ABG64522;  
XX  
DT 27-ANG-2002 (first entry)  
XX  
DE Human albumin fusion protein #1197.  
XX  
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW hematopoietic disorder; neural disorder; connective disorder;  
KW cytostatic; antifertility; antiinflammatory; anticancer;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
XX osteopathic; antarthritic.  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200177137-A1.  
XX  
PD 18-OCT-2001.  
XX  
PF 12-APR-2001; 2001WO-US11988.  
XX  
PR 12-APR-2000; 2000US-229358P.  
PR 25-APR-2000; 2000US-199384P.  
PR 21-DEC-2000; 2000US-256931P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Haseltine WA;  
XX  
DR WPI: 2002-010886/01.  
XX  
PT New fusion protein for treating disease e.g. diabetes comprises an  
XX albumin fused to a therapeutic protein -  
XX  
PS Claim 1; Page 1288-1291; 2102pp; English.  
XX  
CC The present invention relates to albumin fusion proteins comprising a  
CC therapeutic protein X and human albumin (HA, also known as human serum  
CC albumin, HSA). The proteins are useful for treating a disease or  
CC disorder that may be modulated by therapeutic protein X. The albumin  
CC extends the shelf-life of protein X, and may increase its biological  
CC in vitro/in vivo activity. The protein is useful for treating and  
CC diagnosing disorders such as cancer, reproductive disorders, digestive  
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
CC (e.g. diabetes), hematopoietic disorders, neural disorders  
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders  
CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin  
CC fusion proteins of the invention.  
XX  
SO Sequence 832 AA:  
Query Match 38.5%; Score 1240; DB 23; Length 832;  
Best Local Similarity 45.4%; Pred. No. 4,3e-116;



•  
•  
•  
•

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:42:14 ; Search time 26 Seconds  
(without alignments)  
1077.997 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219  
Sequence: 1 MAPHAWVLLARLWGIG.....PCHLSSNALPDSNPESV 596

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2720	84.5	834	1	SM4C_MOUSE
2	1240	38.5	832	1	SM4B_HUMAN
3	1225.5	38.1	838	1	SM4G_HUMAN
4	1223.5	38.0	837	1	SM4G_MOUSE
5	1222.5	38.0	861	1	SM4D_MOUSE
6	1219.5	37.9	862	1	SM4D_HUMAN
7	1149	35.7	782	1	SM4B_MOUSE
8	1146.5	35.6	766	1	SM27_BRARE
9	1029.5	32.0	761	1	SM4A_HUMAN
10	1008.5	31.3	776	1	SM4F_RAT
11	1005.5	31.2	770	1	SM4F_HUMAN
12	1000.5	31.1	777	1	SM4F_MOUSE
13	992.5	30.8	760	1	SM4A_MOUSE
14	937.5	29.1	772	1	SM3A_MOUSE
15	937.5	29.1	772	1	SM3A_RAT
16	934.5	29.0	772	1	SM3A_CHICK
17	932.5	29.0	771	1	SM3A_HUMAN
18	913	28.4	778	1	SM1B_BRARE
19	902.5	28.0	761	1	SM3D_CHICK
20	895.5	27.8	777	1	SM3D_HUMAN
21	892	27.7	749	1	SM3B_HUMAN
22	890	27.6	860	1	SM1A_BRARE
23	872.5	27.1	748	1	SM3B_MOUSE
24	844.5	26.2	775	1	SM3E_HUMAN
25	841.5	26.1	775	1	SM3E_MOUSE
26	837	26.0	764	1	SM22_BRARE
27	828	25.7	785	1	SM3F_MOUSE
28	827	25.7	785	1	SM3F_HUMAN
29	804.5	25.0	785	1	SM3E_CHICK
30	793.5	24.7	751	1	SM3C_CHICK
31	779	24.2	751	1	SM3C_HUMAN
32	774	24.0	751	1	SM3C_MOUSE
33	739	23.0	712	1	SM1A_CAEL

34	730	22.7	771	1	SM1A_DROME	Q24322 drosophila
35	729.5	22.7	1074	1	SM5A_HUMAN	Q13591 homo sapien
36	729	22.6	1077	1	SM5A_MOUSE	Q62217 homo sapien
37	726.5	22.4	930	1	SM6C_HUMAN	Q9h3c2 homo sapien
38	719.5	22.4	295	1	SM4D_CHICK	Q90665 gallus galli
39	708.5	22.0	730	1	SM1A_SCHAM	Q26473 schistocerc
40	708	22.0	1093	1	SM5B_MOUSE	Q60519 mus musculu
41	706.5	21.9	931	1	SM6C_MOUSE	Q9wtm3 mus musculu
42	704.5	21.9	960	1	SM6C_RAT	Q9wtl3 rattus norv
43	704	21.9	697	1	SM2A_SCHGR	Q9xcz8 schistocerc
44	686	21.3	1030	1	SM6A_HUMAN	Q9h266 homo sapien
45	674	20.9	888	1	SM6A_MOUSE	Q35464 mus musculu

ALIGNMENTS

RESULT 1  
SM4C\_MOUSE STANDARD: PRT: 834 AA.  
ID SM4C\_MOUSE  
AC Q64151;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Semaphorin 4C precursor (Semaphorin I) (Sema I) (Semaphorin C-like 1)  
DE (M-Sema F).  
GN SEMA4C OR SEMA1 OR SEMA1L.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Neonatal brain;  
RX MEDLINE=95385809; PubMed=7656991;  
RA Inagaki S., Furuyama T., Iwanashi Y.;  
RT Identification of a member of mouse semaphorin family.\*;  
RL FEBS Lett. 370:269-272(1995).  
RN [2]  
RP INTERACTION WITH GIPC.  
RX MEDLINE=99253973; PubMed=10318831;  
RA Wang L.-H., Kald R.G., Stiltmayer S.M.;  
RT "A PDZ protein regulates the distribution of the transmembrane  
semaphorin, M-SemF.\*";  
RL J. Biol. Chem. 274:14137-14146(1999).  
CC -1- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: EXPRESSED WIDELY IN THE NERVOUS TISSUES DURING  
DEVELOPMENT. PREDOMINANTLY IN THE DEVELOPING BRAIN AND SPINAL  
CORD.  
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC -1- SIMILARITY: Contains 1 Sema domain.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL: S79463; AAB35184.1; -  
DR PIR: S66498; S66498.  
DR MGD: MGI:109252; Sema4C.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003659; Plexin-like.  
DR InterPro: IPR002165; Plexin\_repeat.  
DR InterPro: IPR001627; Sema.  
DR Pfam: PF01437; PSI; 1.  
DR Pfam: PF01403; Sema; 1.

DR SMART; SM00409; IG; 1.  
 DR SMART; SM00423; PSI; 1.  
 DR SMART; SM00630; Sema; 1.  
 DR Signal; PS00835; IG-LIKE; 1.  
 KM Signal; Transmembrane; Immunoglobulin domain; Multigene family;  
 KM Neurogenesis; Developmental protein; Glycoprotein.  
 FT SIGNAL 1 20  
 FT CHAIN 1 834  
 FT DOMAIN 21 664  
 FT TRANSMEM 665 685  
 FT DOMAIN 686 834  
 FT DOMAIN 53 481  
 FT DOMAIN 489 552  
 FT DOMAIN 557 645  
 FT DOMAIN 678 687  
 FT DOMAIN 709 774  
 FT DOMAIN 831 834  
 FT DISULFID 578 628  
 FT CARBOHYD 106 106  
 FT CARBOHYD 121 121  
 FT CARBOHYD 310 310  
 FT CARBOHYD 419 419  
 FT CARBOHYD 522 522  
 FT CARBOHYD 565 565  
 SQ SEQUENCE 834 AA; 92556 MW; 6668B5B5F71482D CRC64;

Query Match 84.5%; Score 2720; DB 1; Length 834;  
 Best Local Similarity 83.1%; Pred. No. 6,4e-221;  
 Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;

1 MAPHNAVMLLARGLGIGAEVMMNIVPRKTVSSGELATVYRRSQIOGDFLTTLTE 60  
 1 MAPHNAVMLLARGLGIGAEVMMNIVPRKTVSSGELATVYRRSQIOGDFLTTLTE 60  
 61 PTGLLYGARREALFAFSMEALDQALISWEAVEKEKTECIQKGNQTECFNFIETLOPY 120  
 61 HSGLLYGARREALFAFSMEALDQALISWEAVEKEKTECIQKGNQTECFNFIETLOPY 120  
 121 NASHLYVCGTAFQPKCTYVNNLTFTLEGEFEDGKCKCPYPAPAGHAGLYDGLYSAT 180  
 121 NASHLYVCGTAFQPKCTYVNNLTFTLEGEFEDGKCKCPYPAPAGHAGLYDGLYSAT 180  
 121 NSSHLYVCGTAFQPKCTYVNNLTFTLEGEFEDGKCKCPYPAPAGHAGLYDGLYSAT 180  
 181 LNPFGLTEPIILRNKGPPIHSMKTEFLAFNLNPHFVGSAYVESYSGFTGDDDKYFEFR 240  
 181 LNPFGLTEPIILRNKGPPIHSMKTEFLAFNLNPHFVGSAYVESYSGFTGDDDKYFEFR 240  
 181 LNPFGLTEPIILRNKGPPIHSMKTEFLAFNLNPHFVGSAYVESYSGFTGDDDKYFEFR 240  
 181 LNPFGLTEPIILRNKGPPIHSMKTEFLAFNLNPHFVGSAYVESYSGFTGDDDKYFEFR 240  
 241 ERAVESDCYAEQVAVARAVKCGMDGAGARTLOKRTTFLKALVCSAPDMKTYFNOLAV 300  
 241 ERAVESDCYAEQVAVARAVKCGMDGAGARTLOKRTTFLKALVCSAPDMKTYFNOLAV 300  
 301 HTLQDTSMHNTTFFGVFOAQOMDYLALCEYOLEIQRVEGPKYKEYHEEAQMDRYTD 360  
 301 HTLQDTSMHNTTFFGVFOAQOMDYLALCEYOLEIQRVEGPKYKEYHEEAQMDRYTD 360  
 301 HTLRASVHNTTFFGVFOAQOMDYLALCEYOLEIQRVEGPKYKEYHEEAQMDRYTD 360  
 361 PVPSPRPGSCINNMHRHGYTSSLPLPNILNFKHPLMEQVCPKMSRPILLYKGTNF 420  
 361 PVPSPRPGSCINNMHRHGYTSSLPLPNILNFKHPLMEQVCPKMSRPILLYKGTNF 420  
 361 PVPSPRPGSCINNMHRHGYTSSLPLPNILNFKHPLMEQVCPKMSRPILLYKGTNF 420  
 421 THLVADRYTGLDGAATYTLFICTGDMWLKAVSLGSPVHLIELDLEDEPARSLVLSQS 480  
 421 THLVADRYTGLDGAATYTLFICTGDMWLKAVSLGSPVHLIELDLEDEPARSLVLSQS 480  
 421 THLVADRYTGLDGAATYTLFICTGDMWLKAVSLGSPVHLIELDLEDEPARSLVLSQS 480  
 481 KKLFLAGRSQSLVOLPVADCIYRSCADCVIARDPYCAWVTSRCVA-VGHSGLSLIQ 539  
 481 KKLFLAGRSQSLVOLPVADCIYRSCADCVIARDPYCAWVTSRCVA-VGHSGLSLIQ 539  
 481 KKLFLAGRSQSLVOLPVADCIYRSCADCVIARDPYCAWVTSRCVA-VGHSGLSLIQ 539  
 540 HVMSTDTGICNLRG-SKIOSGPKRNTTYVAGTDLVPCCHLSNLA-----LPD 588  
 541 HVMSTDTGICNLRG-SKIOSGPKRNTTYVAGTDLVPCCHLSNLA-----LPD 588  
 589 SNP 591  
 599 BGP 601

RESULT 2  
 ID SM4B\_HUMAN STANDARD: PRT: 832 AA.  
 AC Q9NPR2; Q9NPR9; Q9NPR5; Q9NPR8; Q9NPR1; Q9NPR6; Q9NPR0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Semaphorin 4B precursor.  
 GN SEM4B OR KIAA1745.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RX MEDLINE-21082932; PubMed-11214970;  
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,  
 RT Prediction of the coding sequences of unidentified human genes. XIX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RN DNA Res. 7:347-355(2000).  
 RN [2]  
 RP SEQUENCE OF 160-832 FROM N.A., AND VARIANT ALA-792.  
 RA Carim L., Estivill X., Escarceller M., Sumoy L.;  
 RN Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 358-832 FROM N.A.  
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Iyama Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Okeyashi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 RT "NEO human cDNA sequencing project."  
 RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 303-832 FROM N.A., AND VARIANT ALA-792.  
 RC TISSUE-Colon, and Pancreas;  
 RX MEDLINE-22389257; PubMed-12477932;  
 RA Strusberg R.L., Felngold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Hsieh F.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mollighy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywicki M.T., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.W., Maira M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO  
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -1- SIMILARITY: Contains 1 Sema domain.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown in position 709  
 CC onward due to a frameshift.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AB051532; BAB21836.1; -  
 DR EMBL; AL390080; CAB98204.1; -  
 DR EMBL; AL390081; CAB98205.1; -  
 DR EMBL; AL390082; CAB98206.1; -  
 DR EMBL; AK026133; BAB15372.1; ALT\_FRAME.  
 DR EMBL; BC010701; AAH10701.1; ALT\_INIT.  
 DR EMBL; BC017658; AAH17658.1; -  
 DR Genew; HGNC:10730; SEMA4B.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PF01437; PSI; 1.  
 DR Pfam; PF01403; Sema; 1.  
 DR SMART; SM00423; PSI; 1.  
 DR SMART; SM00630; Sema; 1.  
 KW Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;  
 KW Developmental; protein; Glycoprotein; signal; Polymorphism.  
 FT SIGNAL 1 38  
 FT CHAIN 39 832  
 FT DOMAIN 39 712  
 FT TRANSMEM 713 733  
 FT DOMAIN 734 832  
 FT DOMAIN 65 502  
 FT DOMAIN 520 574  
 FT DOMAIN 599 658  
 FT DOMAIN 753 776  
 FT DISULFID 606 651  
 FT CARBOHYD 64 64  
 FT CARBOHYD 91 91  
 FT CARBOHYD 160 160  
 FT CARBOHYD 405 405  
 FT CARBOHYD 520 520  
 FT CARBOHYD 625 625  
 FT VARIANT 792 792  
 FT CONFLICT 160 160  
 FT CONFLICT 565 565  
 SQ SEQUENCE 832 AA; 92192 MW; 29D58C1DD55E1C6B CRC64;  
 Query Match 38.5% Score 1240; DB 1; Length 832;  
 Best Local Similarity 45.4%; Pred. No. 2,6e-96;  
 Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

QY 25 WNLVPRVTVSSGELATVVRPSGICDDELITLTPTGLTYGAREALPFSMEALEIQ 84  
 DB 37 WALSPRISLPLSGSEERPLRFREARHISNYTALISRDGRITLYGAREALFALSINLSFLP 96  
 QY 85 GA-----ISWEAPVKEKTECIQNGKNNOTECFNFRIFLOPYNAVSLYGYTAQPKCTYV 140  
 DB 97 GGEVQELLMKGADEAKKQCKSPKGDPOKCONYKILPLPSGSHLFTCGTAASFPMCTYI 156  
 QY 141 NMLTFLIEHG-----FEDGKGCPCYPDPAGHAGLVDGELYSATLNNFLTEBILNNM 195  
 DB 157 NMENFTLARDKGNVLEDEKGRCPDPNFKSTALVVDGELYTGVTSFQGNDAISMSQ 216  
 QY 196 GPHSMKTEYLAFYLNPRHYGSAVYVESGSPFGDDDKYFFERRAREVSDCAEOVYA 255  
 DB 217 SLRPT-FTESSLNLDPAFVAVASAYIPESLGSLOGDDDKYFFESSEGOEFEPFENTVS 275  
 QY 256 RVAVACGKDGAGATLQKRTTFKARLACAPRNQULFYNLOLQAMHTLQDT--SMHNTF 313  
 DB 276 RIRACICGDEGGEVYLDQKRTSFLKADLLCSRPDGGPFVNLQVFTLSPSPOMRDLTF 335  
 QY 314 FGVAQAW--GDVTLAICEVLEETQVREGPKYHEEAQKMDRYTDPVSPSPSCI 371  
 DB 336 YGVTSOMHRTGTEGSAVCFYTMKVQVRSGLYKVNRETOQVYTVTHVPTDRPACI 395  
 QY 372 NNHRRRGITSSLEPNNILNFVKKHPLMEQVQPKRSRLLYVKKGNFTHLVADRYTGL 431

DB 396 TNSAREKINSIQLPDRIVNLKDHFLMDQVR---SRMLLQPOARYORAVHVRPGL 452  
 QY 432 DGATVTVLEGTGGMILKAVSLAPVWHLIEQLDFD--EPMSRIVLSQSKLLPAGRS 490  
 DB 453 H-HYDVFLEGTGGRKRAKAVSPRRVHIIIEQLIFSSGOVQVLLDTHRGLLYASHS 511  
 QY 491 QLVOLPVADCIKTRSCADCVLARDPYCAMSVNTRCAVAGHPGSLT-----TQHVMTSDT 546  
 DB 512 GVGVVPVAMNCSLYRSCGCLLARDPYCAMS--GSSCKRIVSLQPOLATRPIDIDEGASA 569  
 QY 547 SGICN-----LGSKIQSGPAP-KNITYVAGDVLVPCHLSSNLA 585  
 DB 570 KDLCSASSVSPSPVPTGKPCDQVOPOPNVTNVLACPLLSNLA 613  
 RESULT 3  
 ID SM4G\_HUMAN STANDARD; PRT; 838 AA.  
 AC 09NTN9; 09HCF3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Semaphorin 4G precursor.  
 GN SEMA4G OR KIAA1619.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Heath P.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20450683; PubMed=10997877;  
 RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;  
 RT \*Prediction of the coding sequences of unidentified human genes.  
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which  
 RT code for large proteins in vitro.\*;  
 RL DNA Res. 7:273-281(2000).  
 CC CC -1- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE (BY SIMILARITY).  
 CC CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC CC -1- SIMILARITY: Contains 1 Sema domain.  
 CC CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

FT DOMAIN 18 675 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 676 636 POTENTIAL.  
 FT DOMAIN 697 838 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 56 489 SEMA.  
 FT DOMAIN 507 558 PSI.  
 FT DOMAIN 567 649 IG-LIKE C2-TYPE.  
 FT DOMAIN 565 568 POLY-PRO.  
 FT DOMAIN 763 774 POLY-PRO.  
 FT DISULFID 384 632 BY SIMILARITY.  
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 538 538 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 543 543 R -> RSQSR (IN REF. 2).  
 SQ SEQUENCE 838 AA; 91496 MW; 98281AEE861F245 CRC64;

Query Match 38.1%; Score 1225.5; DB 1; Length 838;  
 Best Local Similarity 46.4%; Pred. No. 4,3e-95;  
 Matches 267; Conservative 79; Mismatches 194; Indels 35; Gaps 15;

29 PKRTVSSGELATVVRFSQGTQIDFLTLTEPTGLLYGAREALFAFSEMALELQGA-- 86  
 34 PRRTIPELST--RHFKGQANTSTLLERASRLVGARGALELSANDIG-DGAIK 90  
 87 -ISMENPVEKTECIQKGNKNOTECNFRFLQPYNASHLYVGYAFOPKCTYVNMILTF 145  
 91 ELMHESAPMSQKCKQKNNOTECFNHFRFLQRLNSTHLYAGCAHAFQPLCAIDAEMAF 150  
 146 TLEHGEFEDKGCPCYDPKAGHAGLLVDGELYSATLNNLGEPIILRMNGHHHSKTEY 205  
 151 TLP-1SFEHGKCKCPYDPKAGFTGLIDGLYTAATRYERST-PDIRSRHP-HSLRTEE 207  
 206 LAF-WLNEPFGVSAVYVESVSGFTGDDDKYFFERERAVE-----SDCYAEQVARYA 258  
 208 TFMHMLNDAEFYFVLYRESKASAVGDDKYYFFTERTEGSGSFGTSRSHVARA 267  
 259 RCKCDNGCARTLQKRTTFLKARLACAPNNQLYFNQLOAMHTLODTSWNTTFGVF- 317  
 268 RCKCDGLGCKTLQKRTTFLKARLACAPNNQLYFNQLOAMHTLODTSWNTTFGVF- 323  
 318 -QAGMDMYLAICEXQLEIQREVEGPKYHEEAKMDRYTDVPSRPGSCINNMHR 376  
 324 LSTOKKTLEBAICRYDLAEIOAVAGPRMETQDSRRKRGEGVPEPRPGSCITDSLR 383  
 377 RHGYTSSLELDPNLIINFKKHPLMEEQVGRMSRPLVKKGTNTHLVADRYTGIDGATY 436  
 384 SQGYNSODLPSLVLDYFKLHPARVYPTGRGRLLLKRNIRYTHLGTPTTPAGPTY 443  
 437 TYLFTGTDGMLLKAVSLCPWHLLEELQF-DQPRMSLVLSQSKLLFAGSRQLOVL 495  
 444 DLLEFTADGWIHKRVVLSGCHHILETQVRESSEVELYLSLQHSLYVAPGAVIQL 503  
 496 PVADCIKRYSCADYIARDPYCANSVNTSRCAVAVGHRG-SLLIOHVTSDSGICNLRG 554  
 504 PLSSSRKRSCTCDLARDPYCGMDPGTHACAAATTINRTALIODIERGN-----RG 556  
 555 --SKIQSGXP--KNITVYACTDLVLPCHLSNLA 585  
 557 CESSHDTGPPPLKRTSVLRGDDVLLPCDQPSNLA 591

RESULT 4  
 SMAG\_MOUSE STANDARD; PRT; 837 AA.  
 AC 09MUT7;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Semaphorin 4G precursor.  
 GN SEMA4G.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_Taxid-10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J;  
 RX MEDLINE-99425180; PubMed-10495281;  
 RA Li H., Wu D.K., Sullivan S.L.;  
 RT Characterization and expression of sema4g, a novel member of the  
 RL semaphorin gene family.  
 Mech. Dev. 87:169-173(1999).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN AXON GUIDANCE.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, AND SEVERAL SENSORY ORGANS  
 CC AS WELL AS SPECIFIC POPULATIONS OF PROJECTION NEURONS.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -1- SIMILARITY: Contains 1 Sema domain.  
 CC -1- SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL; AF134918; AAD30541.1; -  
 DR MGD: MGI:1347047; Sema4g  
 DR InterPro: IPR007110; Ig\_1like.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003659; Plexin-like.  
 DR InterPro: IPR002165; Plexin\_repeat.  
 DR InterPro: IPR001627; Sema.  
 DR Pfam: PF01437; PSI; 1.  
 DR Pfam: PF01403; Sema; 1.  
 DR SMART: SM00409; Ig; 1.  
 DR SMART: SM00423; PSI; 1.  
 DR SMART: SM00630; Sema; 1.  
 DR PROSITE: PS50835; Ig\_Like; 1.  
 KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;  
 KW Neurogenesis; Developmental protein; Glycoprotein.  
 FT SIGNAL 1 17  
 FT CHAIN 18 837  
 FT DOMAIN 18 673  
 FT TRANSMEM 674 694  
 FT DOMAIN 695 837  
 FT DOMAIN 56 487  
 FT DOMAIN 505 558  
 FT DOMAIN 565 647  
 FT DOMAIN 563 566  
 FT DOMAIN 762 773  
 FT DISULFID 582 630  
 FT CARBOHYD 55 55  
 FT CARBOHYD 111 111  
 FT CARBOHYD 126 126  
 FT CARBOHYD 386 386  
 FT CARBOHYD 540 540  
 FT CARBOHYD 538 538  
 SQ SEQUENCE 837 AA; 92378 MW; 3C6E9C98A545EB1 CRC64;

Query Match 38.0%; Score 1223.5; DB 1; Length 837;  
 Best Local Similarity 46.9%; Pred. No. 6,3e-95;  
 Matches 268; Conservative 82; Mismatches 192; Indels 29; Gaps 16;

29 PKRTVSSGELATVVRFSQGTQIDFLTLTEPTGLLYGAREALFAFSEMALELQGA 86  
 34 PRRTIPELSTQIRHFKGCT--QNTSTLLERASRLVGARGALELSANDIRORTIKE 91  
 87 ISMENPVEKTECIQKGNKNOTECNFRFLQPYNASHLYVGYAFOPKCTYVNMILTF 146

Dh	92	IHMVASPEMOSKCHQKGNNOETECFNHFRLORLNATHEFYACGTHAFQPLCAADIAETFI	151
Qy	147	LEHEBEDGKCKCYDPAKAGLILVDELYSATLNNPLGTEPILIRMGCHHSKITYL	206
Dh	152	LP-NSFEFGKCKCYDPAKAGLILVDELYSATLNNPLGTEPILIRMGCHHSKITYL	208
Qy	207	AF-PLNEHFVGSAYVPEVSGSFTGGDDKVFEEFPERAVESDYAE---QVAVARAVC	261
Dh	209	PMHMLNDHEFVSFLVARESKTSVAGDDDKITFFPEREBSGSSFTQSSSRVAVARVC	268
Qy	262	KDGMGARTLQRRKTTPLKARILACSAPNMOUYFNOLQAMHTLQ-DTSMHNTTFEGVF--Q	318
Dh	269	KGDIGKRIKQKMTSFLKARILCHIPQE-----TLRGVGLNDTSSH-THFVAVFTLT	323
Qy	319	AQMGDMYISAICEFQLEBTOVAFBGPKEVHEAEAKMDRYDPPSPPGSCINNMRHH	378
Dh	324	TQMTTLEASACRYDLAEIQAVFTGCPMEYDGDGARMRGVEBGVDEPFGSCITDSLSR	383
Qy	379	GYSLSLEPDMLNPMVKHPRKPMEOQVGRMSRPLLYKGTNFTLVDNRVGLDQATVY	438
Dh	384	GYSNSQDLPSTLVDPVKLHPMARVPVTRGRPLLNKVNARTLTLGTHVSTPAPRYDL	443
Qy	439	LFITGCGMLIKAVSLGFWHLIEQLF-DQEPKRSVLSQSKILFAGSRSLQVLPV	497
Dh	444	LFGLTADCMHKRAYVLGSGMHIIIEIQVFRPQSDNLVISMQSHLYGGAASGLQFPL	503
Qy	498	ADCIKTRSCADCVLARDFPCAMSVNTSRC-VAVGHFSGSLLIQHYMTSDTSGICNLGRSK	556
Dh	504	SSGRGYGVCYDCLLARDPYCGMDSIHACMVATVYANTELIIQDI---EGNRCCEBSR	559
Qy	557	IQSGXP--KNITVYAGTDLVLPCHLSSNLA	585
Dh	560	DAGPPEPLKTRSVLRGDDVLLPCDQPSNLA	589
RESULT 5			
SM4D_MOUSE			
AC	009126;		
AD	30-MAY-2000 (Rel. 39, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Semaphorin 4D precursor (Semaphorin J) (Sema J) (Semaphorin C-like 2)		
DE	(M-Sema G).		
GN	SEMA4D OR SEMAJ OR SEMACL2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6; TISSUE=Brain;		
RX	MEDLINE=97125976; PubMed=896198;		
RA	Furuyama T., Inagaki S., Kosugi A., Noda S., Saitoh S.-I., Ogata M.,		
RA	Iwahashi Y., Miyazaki N., Hamanaka T., Toyama M.;		
RT	Identification of a novel transmembrane semaphorin expressed on		
RT	lymphocytes.;		
RL	J. Biol. Chem. 271:33376-33381(1996).		
CC	- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL		
CC	AS IN THE NERVOUS SYSTEM.		
CC	- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN LYMPHOID TISSUES,		
CC	ESPECIALLY IN THE THYMUS, AS WELL AS IN THE NERVOUS TISSUES.		
CC	- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.		
CC	- SIMILARITY: Contains 1 Sema domain.		
CC	- SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	at the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a> ).		

or send an email to [license@slsb.slb.ch](mailto:license@slsb.slb.ch).

CC										
DR	EMBL	U69535	AAC52964.1	-						
DR	MGI	MGI:109244	Sema4d							
DR	InterPro	IPR007110	Ig-like							
DR	InterPro	IPR003598	Ig_C2							
DR	InterPro	IPR003006	Ig_MHC							
DR	InterPro	IPR003659	Plexin-like							
DR	InterPro	IPR002165	Plexin_repeat							
DR	Pfam	PF00047	Ig_1							
DR	Pfam	PF01437	PSI_1							
DR	Pfam	PF01403	Sema_1							
DR	SMART	SM00408	IGC2_1							
DR	SMART	SM00423	PSI_1							
DR	SMART	SM00630	Sema_1							
DR	PROSITE	PS50835	IG_LIKE_1							
KM	Signal	Transmembrane; Immunoglobulin domain; Multigene family;								
FT	Neurogenesis	Developmental protein; Glycoprotein.								
FT	SIGNAL	1	23							POTENTIAL.
FT	CHAIN	24	861							SEMAPHORIN 4D.
FT	DOMAIN	24	733							EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	734	754							POTENTIAL.
FT	DOMAIN	755	861							CYTOSOLSMIC (POTENTIAL).
FT	DOMAIN	50	482							SEMA.
FT	DOMAIN	502	551							PSI.
FT	DOMAIN	555	636							IG-LIKE C2-TYPE.
FT	DISULFID	576	624							BY SIMILARITY.
FT	CARBOHYD	49	49							N-LINKED (GLCNAC. . .)
FT	CARBOHYD	77	77							(POTENTIAL).
FT	CARBOHYD	139	139							(POTENTIAL).
FT	CARBOHYD	191	191							(POTENTIAL).
FT	CARBOHYD	379	379							(POTENTIAL).
FT	CARBOHYD	419	419							(POTENTIAL).
FT	CARBOHYD	613	613							(POTENTIAL).
FT	CARBOHYD	632	632							(POTENTIAL).
SO	SEQUENCE	861 AA	95714 MM	533CD6D271A6D79B	CRC64					
Query Match										
Best Local Similarity 38.0%; Score 1222.5; DB 1; Length 861;										
Matches 268; Conservative 82; Mismatches 197; Indels 21; Gaps 12;										
QY	28	VPRKTVSSGELATVVRPSPOTGIDFTLTLTTEPTGLLYVGAREALFAF-SMEALLELOGA	86							
DB	27	VPRLTWHEGEVGLV--QFHKRGLINYSALLMSSEDKDLYVGAREAVAVNALNTSEKQHE	84							
QY	87	ISMADPVPKTECTLOKQKKNNTDECFNTRFLQPNASHLYYCCTYATQPKCTYNNMLFT	146							
DB	85	VYMWVSDEKSKCKEKGSKSTECTELINTLVLTLPSSSTLYCGTNAFOPTDHLNLSFK	144							
QY	147	LEHEPEDEGKCKCYDPAKGAAGLLVNGELYSATLNLFGLTEPTILRNMGPHHSKTEYL	206							
DB	145	F-LQKSEDEGKCRCPFDPAHASTVYMGVGEIYSGSYNPLSGEPILISRR-SHSPLRTEYA	202							
QY	207	AFWLNDEPHVGSAYVPSVGSFTGDDKVFYFFERAEVSDCYAEQVAVARVAVCKDGMG	266							
DB	203	IPMNLSEFVADVIQKSPDGEDEDDKVFYFFFEVSGEYEFVFKLMPRAVAVCKDGMG	262							
QY	267	GARTLOKRRKTFELARLACASPNMQLFNLQALAHNTLQDTSWNTHTFFGVQAOAGDMYL	326							
DB	263	GLRTLOKRRKTFELARLACASPNMQLFNLQALAHNTLQDTSWNTHTFFGVQAOAGDMYL	322							
QY	327	SAICEYOLSEETIORVF-EGPYKE---YHEAOQKMYRDPVSPSPRGSCINNMHRRHGYTS	382							
DB	323	SAVCAVILATYEAFAVFSRQKTYQSGATVEQSSHKMYRINQVPTPPGACIDSEARAANTYS	382							
QY	383	SLLELDNLNLFVKKHPLIMEEQVGRWRSRPLLYVKKGTFTFHLVADRVTLGAGATYVLFIG	442							
DB	383	SLNLPDKTQLQVKKHPLIMDSVTFIDNRPKIKKDVNTQLYVDRITQALDSTFIDVAFIS	442							
QY	443	TGDCGMLKAVSLGFWHLLIEELQLP-DOEPMRSVLVSQSK--KILFAGSNQVOLQPVAD	499							
DB	443	TDRALTKRAVILTEEVHVIETQLTFRPEFVLLILSSKGRKRVYAGSNQVQAPLAF	502							

QY 500 C1KYN3CADCVLARDPYCAMSVNTSRCAVGHFSSL-L10HVMTSITSGICNLSKRI 557  
DR SMART; SM00630; Sema; 1.  
DR PROSITE; PS50835; IG-Like; 1.  
KW Signal: Transmembrane; Immunoglobulin domain; Multigene family;  
KW Neurogenesis; Developmental protein; glycoprotein.  
DB 503 CCKHSCEDCVLARDPYCAMSPAIKACVTLH0EASRRGMID-MSGDTSS-CLDSKKS 560  
QY 556 QSGPKXKNTTVVAGDVLVPLCHLSSNLA 585  
DB 561 FNOHFFKH-----GGTAEIKCFKSNLA 583  
RESULT 6  
SMAD\_HUMAN STANDARD; PRT; 862 AA.  
ID SMAD\_HUMAN 092854;  
AC 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Semaphorin 4D precursor (Leukocyte activation antigen CD100) (BB18)  
DE (A8) (GR3).  
GN SEMA4D OR CD100.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=T-cell;  
RX MEDLINE=97030273; PubMed=8876214;  
RA Hall K.T., Boumsell L., Schultze J.L., Boussioutis V.A., Dorfman D.M.,  
Cardoso A.A., Benussan A., Nadler L.M., Freeman G.J.;  
RT "Human CD100, a novel leukocyte semaphorin that promotes B-cell  
aggregation and differentiation." J. Biol. Chem. 268:11780-11785 (1996).  
RT Proc. Natl. Acad. Sci. U.S.A. 93:11780-11785 (1996).  
CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL  
AS IN THE NERVOUS SYSTEM. INDUCES B CELLS TO AGGREGATE AND  
IMPROVES THEIR VIABILITY IN VITRO.  
CC -1- SUBUNIT: Homodimer.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN SKELETAL MUSCLE,  
PERIPHERAL BLOOD LYMPHOCYTES, SPLEEN, AND THYMOS AND ALSO  
EXPRESSED AT LOWER LEVELS IN TESTES, BRAIN, KIDNEY, SMALL  
INTESTINE, PROSTATE, HEART, PLACENTA, LUNG, AND PANCREAS BUT NOT  
IN COLON OR LIVER.  
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC -1- SIMILARITY: Contains 1 Sema domain.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD100 entry;  
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd100.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sb.ch/announce/  
or send an email to license@isb-sb.ch).  
CC -----  
DR EMBL; U60800; AAC50810.1; -  
DR Genbank; HGNC:10732; SEMA4D.  
DR MIM; 601866; -  
DR GO; GO:0006916; P:anti-apoptosis; TAS.  
DR GO; GO:0007155; P:cell adhesion; TAS.  
DR GO; GO:0006955; P:immune response; TAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003006; Ig-MHC.  
DR InterPro; IPR003659; Plexin-like.  
DR InterPro; IPR002165; Plexin\_repeat.  
DR InterPro; IPR001627; Sema.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF01437; PSI; 1.  
DR Pfam; PF01403; Sema; 1.  
DR SMART; SM00409; IG; 1.

DR SMART; SM00423; PSI; 1.  
DR SMART; SM00630; Sema; 1.  
DR PROSITE; PS50835; IG-Like; 1.  
KW Signal: Transmembrane; Immunoglobulin domain; Multigene family;  
KW Neurogenesis; Developmental protein; glycoprotein.  
DB 503 CCKHSCEDCVLARDPYCAMSPATCAVALHOTESPSSG-LIQ-MSGDS-VCP 554  
QY 552 LRSGKQSGPKXKNTTVVAGDVLVPLCHLSSNLA 585  
DB 555 DK-----SKGSYRQHFFKHGTT-AELKCSQKSNLA 583  
RESULT 7  
SMAB\_MOUSE STANDARD; PRT; 782 AA.  
ID SMAB\_MOUSE  
Query Match 37.9%; Score 1219.5; DB 1; Length 862;  
Best Local Similarity 46.0%; Pred. No. 1.4e-94;  
Matches 264; Conservative 88; Mismatches 189; Indels 33; Gaps 14;  
QY 28 VPKRYTSSELELTVVRFSGTGTODPLTLTTPRLTVVGAELAF-SHEALDQA 86  
DB 27 IPIITHEHREVLV-QFHEPDLYNSALLSDKDTLIGAEVAVNALISKEQH 84  
QY 87 ISWEAFVEKTECTIOGKNNQTECFNFIPLPYNASHLYVCGTVAOPRCTYVNLTF 146  
DB 85 YVWKVSEDKKAKCAEKGSKQTECLNIVLPDLSMTSLYCGTNAFOPACHLNLTSRK 144  
QY 147 LEHGEFEDCKGKCPYDPAKAGLVDGLYSATLNNFTGTEPIILRNKPHHSKTEYL 206  
DB 145 F-LGKREDDKGRCPDPASHTSVYVADGELYSTGTFSGPILISRN-SSHSPLTEYA 202  
QY 207 AFPLNPHFVSGAYVESGTFGDDDKYFFPFRFNAVSDCAEYVAVARVAVCGDMG 266  
DB 203 IPLNPSFVEADVINKSPDSDEDDRVFFETEVSEYFVRLIPRIAVNCKGDG 262  
QY 267 GARTLQKRTTFLKARLACSAFNMQLYFNQDAMHTLQDTSNHTTFEGVFQAKMDMTL 326  
DB 263 GLRTLOKRTTFLKARLACSAFNMQLYFNQDAMHTLQDTSNHTTFEGVFQAKMDMTL 322  
QY 327 SAICEYQLEIDQRF-EGPYKE---YHEAOKWDRYTDVPSPRPGSCINNMHRRHGYTS 382  
DB 323 SAVCAVNLSTAEVFSHGKRYQWSTVQESHKTVKVRNGVPPRPACIDSEARAANYTS 382  
QY 383 SLELPNLIIEFYKHPHLMEOYGRKSRPLVKGTFNHLVADRTGDAATYTLVFLG 442  
DB 383 SLELPNLIIEFYKHPHLMEOYGRKSRPLVKGTFNHLVADRTGDAATYTLVFLG 442  
QY 443 TGDGWLKAVSLGFWHLTLEIDLF-DOEPMSLVLSQSK-KLTFAGRSQQLVQLPVD 499  
DB 443 TDRGALHKAISLEHNVHIEFTQQLPDPEPVQTLILSSKKNRNFYAGSSNGVQAPLAF 502  
QY 500 C1KYN3CADCVLARDPYCAMSVNTSRCAV-----GHHFSSL10HVMTSITSGICN 551  
DB 503 CCKHSCEDCVLARDPYCAMSPATCAVALHOTESPSSG-LIQ-MSGDS-VCP 554  
QY 552 LRSGKQSGPKXKNTTVVAGDVLVPLCHLSSNLA 585  
DB 555 DK-----SKGSYRQHFFKHGTT-AELKCSQKSNLA 583

AC	062179;2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
GN	SEMA4B OR SEMAC OR SEMC.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-NMRI: TISSUE-Brain;
RX	MEDLINE-95267431; PubMed-7748561;
RA	Pueschel A.W., Adams R.H., Betz H.;
RT	"Murine semaphorin D/collapsin is a member of a diverse gene family
RT	and creates domains inhibitory for axonal extension.";
RL	Neuron 14:941-948(1995).
RN	[2]
RP	INTERACTION WITH GIPC.
RX	MEDLINE-99253973; PubMed-10318831;
RA	Wang L.-H., Kalb R.G., Strittmatter S.M.;
RT	"A PDZ protein regulates the distribution of the transmembrane
RT	semaphorin, M-SemF.";
RL	J. Biol. Chem. 274:14137-14146(1999)
CC	- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC	SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC	- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC	- SUBCELLULAR LOCATION: Type I membrane protein.
CC	- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
CC	LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
CC	MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC	- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC	- SIMILARITY: Contains 1 Sema domain.
CC	- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on les-
CC	use by non-profit institutions as long as its content is in no way com-
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC	or send an email to license@isb-slb.ch).
CC	-----
DR	EMBL: X85992; CAAS9984.1; .
DR	PIR: I48746; I48746.
DR	MGI: MGI:107559; Sema4b.
DR	InterPro: IPRO03659; Plexin-like.
DR	InterPro: IPRO02165; Plexin_repeat.
DR	InterPro: IPRO01627; Sema.
DR	Pfam: PF01437; PSI; 1.
DR	Pfam: PF01403; Sema; 1.
DR	SMART: SM00423; PSI; 1.
DR	SMART: SM00630; Sema; 1.
KM	Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
KM	Developmental protein; Glycoprotein.
FT	NON_TER 1
FT	DOMAIN <1 662 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 663 683 POTENTIAL.
FT	DOMAIN 684 782 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 16 453 SEMA.
FT	DOMAIN 471 525 PSI.
FT	DOMAIN 548 608 IG-LIKE C2-TYPE.
FT	DOMAIN 703 726 PRO-RICH.
FT	DISULFID 555 601 BY SIMILARITY.
FT	CARBOHYD 12 12 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 15 15 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 42 42 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 88 88 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 356 356 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 471 471 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 526 526 N-LINKED (GLCNAG. . . ) (POTENTIAL).
FT	CARBOHYD 574 574 N-LINKED (GLCNAG. . . ) (POTENTIAL).

FT	CARBOHYD	639	639	N-LINKED (GLCNAC	) (POTENTIAL).
SEQ	SEQUENCE	782 AA:	86823 MW:	627A81F0C8B8FAC	CRC64:
	Query Match	35.7%;	Score 1149;	DB 1;	Length 782;
	Best Local Similarity	43.6%;	Pred. No. 1.le-88;		
	Matches 247;	Conservative	89;	Mismatches 201;	Indels 30; Gaps 11
QY	41	VVRNFSQGIODELTLLTLEPTGLLYGAREALFAFSKDALELOGA---	ISWEAPVERK	96	
DB	4	LIRKEAENISNYTALLLSODCKTYLVAAREALFALNSNLSFELPGEGYOEILMSADARK	63		
QY	97	TECIOKGNNOECENFTRLFQPVASHLUYCGTAFQPKCYVMMLFTLEHGE----	151		
DB	64	QOCSTKRGADPKKDCONNTIKILPLNLSHLLTCGTAFAFPLCAYIHAFSTLAQDAGNYI	123		
QY	152	FEDGKCKPPYDPAPAKHAGLLVDGELYSALTLEPTETLLNMGPHSMKTEYAFATYN	211		
DB	124	LEDGCHCPFDNENFSTALVYDDELXTGTVSFGQNDRAISRQSSRPL-KTESSLMLQ	182		
QY	212	EPHFVGSAYVPEBSVSGFTGDDKYVFFPRERAESDCAEOYVAVARVCKDGMGARTL	271		
DB	183	DPAFVASATSPESLSPIGDDDKIYFFSETEGEEFEFENTIVSRVARVCKDGEGERYL	242		
QY	272	QKWTTFELKARLACGAPNMOLYFNQLOAMHTLODT--SKHNTTFEGVFQAGH--GDMYLS	327		
DB	243	QORMTSFLKAOLLCRPPDGEPFNVLQOVETLNPNDQMRKTLISIGVFTSOWHRCSTEGS	302		
QY	328	AICEQTLEIQIOWFEGPYKEYEHEEAKMDRYDTPVPSRPGSCINMHRHAGTISLELP	387		
DB	303	AICVYTMNDVQKAFGLGKLYKKNREIQOQMTTEHVOYTPRPGACITNSAREKINSLOLP	362		
QY	388	DNILNFVKHPLMEQOVRPMRSRPLLYKKGTFHLYVADRYTGLDGAITYVLFITGDMW	447		
DB	363	DRVLNFIHDFHMDQVYR---SRLLTLDPARYQVAVHRYGLH-STYDVLFLGTGDR	418		
QY	448	LLKAYSLEGPWHLIIEHLQFDO-EPKRSIVLYSQSKLLFAGSRQDLQPLVADCIKIRYSC	506		
DB	419	LKRAVTLSSRHIIIEHLQIFPGQGVONMLLDSHGGLLYASSHSGVQVOPVANCSLYPTC	478		
QY	507	ADCVLARPPCYASVNTSRGCVAVGGHFGSLIQIHWTSPTDS-----ICNLRSK--IQ	558		
DB	479	GDCLLARPPCYMTSACRSLASL---YQDPLASRWTDIDEGASVKEILCKNSYARFLV	535		
QY	559	SGPXPKRNTVVAAGTDLVLPCHLSSMLA	585		
DB	536	PKRCKQVOIQENTVNTLACPLSLMLA	562		
RESULT 8					
SM27	BRARE				
ID	SM27_BRARE	STANDARD:	PRT:	766	AA.
AC	09YHX4:				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Semaphorin z7 precursor (Semaphorin 7) (Sema-27).				
GN	SEMA27 OR SEMA7.				
OS	Brachydanio rerio (Zebrafish) (Danio rerio).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;				
OX	NCBI_Taxid=7955;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99060909; PubMed=9667349;				
RA	Halloran M.C., Severance S.M., Yee C.S., Gemza D.L., Kuwada J.Y.;				
RT	"Molecular cloning and expression of two novel zebrafish				
RT	semaphorins."				
RL	Mech. Dev. 76:165-168(1998).				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.				
CC	-1- SIMILARITY: Contains 1 Sema domain.				
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.				

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF073289; AAC7345.1; -  
 CC ZFIN: ZDB-GENE-990715-7; sema7.  
 CC InterPro: IPR002110; Ig-like.  
 CC InterPro: IPR003599; Ig.  
 CC InterPro: IPR003006; Ig\_MHC.  
 CC InterPro: IPR003659; Plexin-like.  
 CC InterPro: IPR002165; Plexin\_repeat.  
 CC InterPro: IPR001627; Sema.  
 CC Pfam: PF00047; Ig\_1.  
 CC Pfam: PF01437; PSI; 1.  
 CC Pfam: PF01403; Sema; 1.  
 CC SMART: SM00409; Ig; 1.  
 CC SMART: SM00423; PSI; 1.  
 CC SMART: SM00630; Sema; 1.  
 CC PROSITE: PS50835; IG\_LIKE; 1.  
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;  
 CC Glycoprotein.  
 CC SIGNAL 1 24 POTENTIAL.  
 CC CHAIN 25 766 SEMAPHORIN 27.  
 CC DOMAIN 25 664 EXTRACELLULAR (POTENTIAL).  
 CC TRAMEM 665 685 POTENTIAL.  
 CC DOMAIN 686 786 CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 33 483 SEMA.  
 CC DOMAIN 501 552 PSI.  
 CC DOMAIN 555 640 IG-LIKE C2-TYPE.  
 CC DISULFID 577 623 BY SIMILARITY.  
 CC CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 766 AA; 85617 MW; 0ACBC693FE7D830C CRC64;  
 CC -----  
 CC Query Match 35.6%; Score 1146.5; DB 1; Length 766;  
 CC Best Local Similarity 42.0%; Pred. No. 17e-88;  
 CC Matches 248; Conservative 101; Mismatches 212; Indels 29; Gaps 14;  
 CC -----  
 CC 7 WLILARLWGLGIGAEVWNNLVPRKTVSSGELATVVRFSQTDITLTTEPTGLLY 66  
 CC 13 WSPAMLVGGLGSLTD---SLPRKTVPIG--SNGGRLEFRGGINVYTMILRLDLNLILI 66  
 CC 67 VCARBALAFRSKEALELQGA-ISMENAPVEKTECIQKGNKNOTECFNFIRIQLPYNASHL 125  
 CC 67 LGARARITFALDDDDITIKKAMLMKWEVTRDQNDCSNKGKDAVDNCKNYIRIHKKNDGRM 126  
 CC 126 YVCGTAYGPKCTYYNNML--FTFLHGEFEQKGCOPYDPAKMGHGLVLDGELYSATLNN 183  
 CC 127 YVCGKARNPCTGSLYADGKLTLEILO-EDTKGKCPDPFRVYSAMVDGAYVATSKNN 185  
 CC 184 FLGTEPILRNKGRPHSMKTEYLALNLPHEFVGSAYPEVSGSTGDDDKYFFERERA 243  
 CC 186 FRGSGPVMMRS--TRESIRTEFTSWLSEPNFIHMAHIFEGSNDGDGDKYLFESERA 243  
 CC 244 VESDCYAAQYVARAVKCKDGMGARTLQKRTTTLKARLACSPANMOLYENQLOAMHTL 303  
 CC 244 VEYESYTYADVSRVAVKCGDLSGORTLQKKTSEFLKARLDQVNTNPLLYQDVHILC 303  
 CC 304 ODTSHNNTFPGVFOAGMDWLSAICGYOLEIORVF-EGYKK-EYHNEAO--FWDRYT 359  
 CC 304 PD-DWTFCTVFAVFPQSDSSQYSAVCSIKEDITIVSKGFKAPNVEVSFVWWMYS 362  
 CC 360 DPVPSRPGSCINNNHRRHGYSTSLPDLNLIYFKKHLMEEOVGPRMSRPLVKKGTN 419  
 CC 363 GELPDRPGACIDNARREKGITKSLFLDKTLQVFKDRPLMDQAVTA--EQLVLVKRGA 420

QY 420 FTHLVADRVYTGDAATVYVLTIGTGDGWLKAVSLGAPVWHLIELQLFD-QEPMNSLYLS 478  
 DB 421 FRIIVYATATATNISHQWYFTGSGSVLKAVYNNGEVIMEEIOLEDPSPKIRLRLS 480  
 QY 479 OSKTLFNGSRQGLVLPVADCIKXRSACDVLADPVCAMSVNTRSCVAGVGH---GS 535  
 DB 481 SKRKOLVYGEVGVVSLSTSECGRYOTCLDVLADPHCGMDLDFEHCRTINSIRTRSS 540  
 QY 536 LTIQHWMTSDTSGICNLNGSKIGSGPKXPKNTIVVAGTDLVPCHESSNLA 585  
 DB 541 TVIQSLNGDASQCPALIGSK-----PVNISFYGNNTYKLCGQCYRNSNA 584  
 RESULT 9  
 SM4A\_HUMAN  
 ID SM4A\_HUMAN STANDARD: PRT; 761 AA.  
 AC Q9H3S1; OSMDA9; (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Semaphorin 4A precursor (Semaphorin B) (Sema B).  
 GN SEMA4A OR SEMB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RP SEQUENCE FROM N.A.  
 RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,  
 RA Saito T.,  
 RT "Human semaphorin B.";  
 RT Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=colon;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Scheenen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Yoshitaki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosack S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Small D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO  
 CC SPECIFIC TERRITORIES INACCESSIBLE FOR GROWING AXONS (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AB029394; BAB20087.1; -





```

QY 87 ISHEAPVEKTECTIOKGNKNNOTECENFIRFLOPYNASHLYVGTAFVOPKCTYVNNLFTF 146
DB 105 IIMMPVETIRHONCRKRGK - KEDECHNFIOILAIVNASHLTCGTAFADPCGVIVDSFQ 163
QY 147 LEHGEFEODKGCOPYPAPKAGHGLVDGELYSATLNFLETETIIRNNPHHS -MKEE 205
DB 164 -OVERLESERGCOPPEAPQASAVMAGVLYATVNFLETETIIRAVGRADMTRTET 222
QY 206 LAEMLEPFRVGSAYV -PEVSGETGDDKRYVEFFERERAVESPCYAEQVAVARVYCKGD 264
DB 223 LSSMLNAPAFVAMVLSPEMGEDEGDE -IEFFETETSRVLDYERIVPRARACAD 281
QY 265 MCGARLQKRTTFLKARLACAPKQQLYFNLOAMHTLQDTSWHNT -TFEGVQAOQMD 323
DB 282 LGGKRLQORWTFLEADLDCGPEHGRASGVLOANAEIRPOGATPTFFYGIFSSOMGS 341
QY 324 MYLSACEIOLEIOFEGEPYKEYHEBAOKMDRYD -PVSPRSPSCINNNRRRGYS 382
DB 342 MAISAVCAFRPDIDRVLNGLFRELKHCNDKCLPVNDNEVPDRPECTANNKLOQFQS 401
QY 383 SLELPDNLINFYKHPIMEEOYGRMSRDLVYKGTFTHLVADRYTGLDGYTYLFTG 442
DB 402 SLSLPDRVLTFRDHPDMRPFVADGRDLVTTDTAVLRVVAHRTSLSGKEYDVLVY 461
QY 443 TDDCHLKAVALSGPWHLIEBQLF -DOEPMSLVLSOSKILLFASRSQVLQVLPVADCI 501
DB 462 TEGHILHRAVRIGAOVSLVEDLALFEPQPVESMKLYD -WLVGSHTEYVQVNTSNG 519
QY 502 KYRSCADVLADPYCAMSNTSRVYA -VGGHFGSLIIQHVMTSDPGICNLGSKIQSG 560
DB 520 RIGQSECLIQDPPVCAMSFRLDACVAHNGEHRG -MVGDISSADVSSIC -PKRG 572
QY 561 PAPKNTTVVAGT -DVLVPCHLSS 582
DB 573 EHPVVEFVPAVATGVHVLPCSPSS 596

RESULT 11
SMAF_HUMAN
ID SMAF_HUMAN STANDARD: PRT; 770 AA.
AC 095754; Q9NS35;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4F precursor (Semaphorin W) (Sema W) (Semaphorin M) (Sema
M).
GN SEMA4F OR SEMA W OR SEMA W.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A. (ISOFORM SMALL).
RC TISSUE=Brain;
RX MEDLINE=99162633; PubMed=10051670;
RA Encinas J.A., Kikuchi K., Chedotal A., de Castro F., Goodman C.S.,
RA Kimura T.,
RA "Cloning, expression, and genetic mapping of Sema W, a member of the
RA semaphorin family."
RA Proc. Natl. Acad. Sci. U.S.A. 96:2491-2496(1999).
RN 12
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC Chen F., Do A., Do T., Weisler M., Roe B.A.;
RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC MEDLINE=2154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgore W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Oltersweider B., Obermaier B., Tampe J., Heubner D.,
RA Wandut R., Korn B., Klein M., Poustka A.;

```

```

RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs."
RT Genome Res. 11:422-435(2001).
RN [4]
RA SEQUENCE OF 347-770 FROM N.A.
RA Jang W., Spillson S.V., Hua A., Roe B., Weisler M.H.;
RA "Large-scale comparative sequence analysis of human and mouse genomic
RA DNA in the mnd2 region of mouse chromosome 6 reveals coding regions of
RA three new genes."
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HAS GROWTH CONE COLLAPSE ACTIVITY AGAINST RETINAL
CC GANGLION-CELL AXONS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named Isoforms-2;
CC Name-Long:
CC IsoId=Q95754-1; Sequence=Displayed;
CC Name-Short:
CC IsoId=Q95754-2; Sequence=VSP_006043;
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB022317; BAA75631.1; -
DR EMBL: AC006544; -; NOT_ANNOTATED_CDS.
DR EMBL: AL136552; CAB66487.1; -
DR EMBL: AF053369; AAF80660.1; -
DR Gene: HGNC:10734; SEMA4F.
DR MIM: 603706; -
DR GO: GO:0005887; C: Integral to plasma membrane; TAS.
DR GO: GO:0005624; C: membrane fraction; TAS.
DR GO: GO:0007411; P: axon guidance; TAS.
DR GO: GO:0007267; P: cell-cell signalling; TAS.
DR GO: GO:0007399; P: neurogenesis; TAS.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin_repeat.
DR Pfam: PF01437; PSI; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; Glycoprotein;
KW Alternative splicing.
FT SIGNAL 1 34
FT CHAIN 35 770
FT DOMAIN 35 659
FT TRANSMEM 660 680
FT DOMAIN 681 770
FT DOMAIN 65 496
FT DOMAIN 512 563
FT DOMAIN 580 635
FT DISULFID 587 628
FT CARBOHYD 64 64
FT CARBOHYD 133 133
FT CARBOHYD 509 509
FT VARSPLIC 120 274
FT CONFLICT 533 533
FT SEQUENCE 770 AA; 83511 MW; 83511 MW; CFB874B410F0E9C8 CRC64;
Query Match 31.2%; Score 1005.5; DB 1; Length 770;
Best local Similarity 39.4%; Pred. No. 1,3e-76;
Matches 222; Conservative 100; Mismatches 222; Indels 19; Gaps 15;

```

```

OY 28 VPKTVSSGELATVVRPSQGTIDFLTLTEPTGLLYGAREALFAFSME-ALELOGA 86
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 40 VPKTSLPSEADSLTREFAPHTYNSVLVDPAHSHTLYGARDTIFALTEPFSEKRR 99
OY 87 ISWAPVPEKTECIQKGNKNOTECGFNFREFLOPYNASHLYVCGTYAPDCKTYVMKLTFT 146
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 100 IDMWPEAHNRCKRKKR-KEDECHNFVQIILANASHLTCGTFAPKCGVIDVSRFO 158
OY 147 LEHGEFEDGKCKCPYDPKAGHAGLYVDGELYSATLNNFLGTEPTILRNMGPHNS-MKTEY 205
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 159 -OVERLESBGRKCFEPFPAORSAAMAGVLAATYKNTLGTPEPTITRAVGAEDMIRDT 217
OY 206 LAFMLNEPHEVGS-AYPEVSGSTGDDDKYFFFRERAVESDCYAEQVAVARVCKGD 264
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 LPSMLNAPAFVAAVALSPAEMGDEGODE-IEFFETFSRAFDSEYERIKVPRVAVACGD 276
OY 265 MGAATLQKRTTFLKARLACSAFNMWLYFNOLQAMHTLQDTSKINT-TFFGVQAOQKD 323
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 277 LGGRKTLQORWTLTKADLLCPGPHGRASSVLDDVAALRELGAGTPIFYGIFSSOWEG 336
OY 324 MYLSAICEYOLEEIORVEGEPYKEYHEBAOKMDRTD-PVPSRPGSCIT-NMHRHNGYT 381
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 337 ATISAVCAFPRQDITRYVNGFRELKHCNGLPVVDVQPRGECITNNMKLRH-FG 395
OY 382 SLELPDNLINLVKKNHMEBOVGPRMSRPLVKKGTFTLVAADRTGLDGAITYVLFT 441
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 396 SLSLDPRLVTLFIRHDLMDRPFADGRLVLTDTAYLRVAVARHTVLSGKEYDVLXL 455
OY 442 GTGGWMLKAVSLGPRWHLIEQLF-DOEPMRSVLVSQSKLLFAGRSQVLQVLPVADC 500
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 456 GTEGHCLRAVRIGQSLVEDLALFPPRPVYENKKLHS--WLLVGRTEVTOVNTNCG 513
OY 501 IKYSCADCVLAPDPYCAVSNTSKVCA-VGSHFESLTIQHWEDTSGICNLGSKIOS 559
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 514 GRGSCSCCLIAOPVYCAMSFRLEDCVNAHGEHRG--LVQIDESADVSLSCKEKEG--- 568
OY 560 GPXRNITVAVAGTDLVLPCHLSS 582
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 569 RPVVEFVAVATAHAYVLPSPSS 591

```

RESULT 12

SM4F\_MOUSE STANDARD: PRY: 777 AA.

```

ID SM4F_MOUSE STANDARD: PRY: 777 AA.
AC 092123; 09R1Y1;
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 15-SEP-2003 (Rel. 42, last annotation update)
DE Semaphorin 4F precursor (Semaphorin W) (Sema W).
GN SEMA4F OR SEMA4F.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c; and 129/SvJ; TISSUE-Brain;
RX MEDLINE=99162633; PubMed=10051670;
RA Enclinas J.A., Kituchi K., Chedotal A., de Castro F., Goodman C.S.,
RA Kimura T.;
RT "Cloning, expression, and genetic mapping of Sema W, a member of the
RT semaphorin family.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2491-2496(1999).
CC -1- FUNCTION: HAS GROWTH CONE COLLAPSE ACTIVITY AGAINST RETINAL
CC GANGLION-CELL AXONS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -1- SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

```

CC EMBL; AB021291; BAA75630.1; -
DB EMBL; AB022316; BAA75634.1; -
DB EMBL; AB022311; BAA75634.1; JOINED.
DB EMBL; AB022312; BAA75634.1; JOINED.
DB EMBL; AB022313; BAA75634.1; JOINED.
DB EMBL; AB022314; BAA75634.1; JOINED.
DB EMBL; AB022315; BAA75634.1; JOINED.
DB MGI; MGI:1340055; Sema4f.
DB InterPro; IPR003659; Plexin-like.
DB InterPro; IPR02165; Plexin_repeat.
DB InterPro; IPR01627; Sema.
DB Pfam; PF01437; PSI; 1.
DB Pfam; PF01403; Sema; 1.
DB SMART; SM00423; PSI; 1.
DB SMART; SM00630; Sema; 1.
KV Signal: Transmembrane; Immunoglobulin domain; Multigene family;
FT Signal: 1
FT CHAIN 41 777
FT DOMAIN 41 667
FT TRANSMEM 668 688
FT DOMAIN 689 777
FT DOMAIN 71 502
FT DOMAIN 518 569
FT DOMAIN 586 641
FT DISULFID 593 634
FT CARBOHYD 70 70
FT CARBOHYD 139 139
FT CARBOHYD 515 515
FT VARIANT 490 490
FT VARIANT 659 659
SQ SEQUENCE 777 AA; 84501 MW; B63F8535856924 CRC64;

```

Query Match 31.1%; Score 1000.5; DB 1; Length 777;

Best Local Similarity 39.2%; Pred. No. 3.4e-76;

Matches 221; Conservative 102; Mismatches 220; Indels 21; Gaps 14;

```

OY 28 VPKTVSSGELATVVRPSQGTIDFLTLTEPTGLLYGAREALFAFSME-ALELOGA 86
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 VPKTSLPSEADSLTREFAPHTYNSVLVDPAHSHTLYGARDTIFALTEPFSEKRR 105
OY 87 ISWAPVPEKTECIQKGNKNOTECGFNFREFLOPYNASHLYVCGTYAPDCKTYVMKLTFT 146
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 106 IDMWPEAHNRCKRKKR-KEDECHNFVQIILANASHLTCGTFAPKCGVIDVSRFO 164
OY 147 LEHGEFEDGKCKCPYDPKAGHAGLYVDGELYSATLNNFLGTEPTILRNMGPHNS-MKTEY 205
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 165 -OVERLESBGRKCFEPFPAORSAAMAGVLAATYKNTLGTPEPTITRAVGAEDMIRDT 223
OY 206 LAFMLNEPHEVGSAYV-PEVSGSTGDDDKYFFFRERAVESDCYAEQVAVARVCKGD 264
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 224 LSSWLNAPAFVAAVALSPAEMGDEGODE-IEFFETFSRAFDSEYERIKVPRVAVACGD 282
OY 265 MGAATLQKRTTFLKARLACSAFNMWLYFNOLQAMHTLQDTSKINT-TFFGVQAOQKD 323
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 283 LGGRKTLQORWTLTKADLLCPGPHGRASSVLDDVAALRELGAGTPIFYGIFSSOWEG 342
OY 324 MYLSAICEYOLEEIORVEGEPYKEYHEBAOKMDRTD-PVPSRPGSCITNNMHRHNGYT 382
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 343 AATISAVCAFPRQDITRYVNGFRELKHCNGLPVVDVQPRGECITNNMKLRH-FG 402
OY 382 SLELPDNLINLVKKNHMEBOVGPRMSRPLVKKGTFTLVAADRTGLDGAITYVLFT 442
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 403 SLSLDPRLVTLFIRHDLMDRPFADGRLVLTDTAYLRVAVARHTVLSGKEYDVLXL 462
OY 442 TDDGWLKAVSLGPRWHLIEQLF-DOEPMRSVLVSQSKLLFAGRSQVLQVLPVADC 501
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 463 TEGHCLRAVRIGQSLVEDLALFPPRPVYENKKLHS--WLLVGRTEVTOVNTNCG 520

```

OY 502 KTRSCADCVLADPVCAMSVNTSRCA-VGHPGSLILIOHWTSDTSGICNKGSIQSG 560  
 DB 521 RQSCSECIADDPVCAMSFRLDACVAHAGEHNG--MVDIESADVSLC-----PKEPG 573  
 OY 561 PPKNTIVVAGT-DVLPCHLSS 582  
 DB 574 EHPVVEPVATVGVHVLPCSPSS 597

RESULT 13  
 SMA\_MOUSE  
 ID SMA\_MOUSE STANDARD: PRT: 760 AA.  
 AC 062178;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Semaphorin 4A precursor (Semaphorin B) (Sema B).  
 GN SEMA4A OR SEMA4B OR SEMB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_Taxid=10090;  
 RX (1)  
 RC SEQUENCE FROM N.A.  
 RX STRAIN-NMRI: TISSUE-brain;  
 RX MEDLINE=95267431; PubMed=7748561;  
 RA Puschel A.W., Adams R.H., Betz H.;  
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family  
 and creates domains inhibitory for axonal extension.";  
 RL Neuron 14:941-948(1995).  
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO  
 SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW  
 LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH  
 MODERATE LEVELS FROM THEN UNTIL BIRTH.  
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -1- SIMILARITY: Contains 1 Sema domain.  
 CC -1- SIMILARITY: Contains 1 Immunoglobulin-like C2-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: X85991; CNA59983.1; -  
 DR PIR: I48745; 148745.  
 DR MGI: 107560; Sema4a.  
 DR InterPro: IPR003659; Plexin-like.  
 DR InterPro: IPR002165; Plexin\_repeat.  
 DR InterPro: IPR001627; Sema.  
 DR Pfam: PF01437; PSI: 1.  
 DR Pfam: PF01403; Sema; 1.  
 DR SMART: SM00423; PSI; 1.  
 DR SMART: SM00630; Sema; 1.  
 DR Signal: Transmembrane; Immunoglobulin domain; Multigene family;  
 KM Neurogenesis6; Developmental protein; glycoprotein.  
 FT SIGNL 1 32  
 FT CHAIN 33 760  
 FT DOMAIN 33 682  
 FT TRANSMEM 683 703  
 FT DOMAIN 704 760  
 FT DOMAIN 64 478  
 FT DOMAIN 496 547  
 FT DOMAIN 572 630  
 FT DOMAIN 609 612  
 FT DISULFID 579 623  
 FT CARBOHYD 120 130  
 FT CARBOHYD 135 135

FT CARBOHYD 496 496 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 606 606 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 760 AA; 83458 MW; 0A4D6B80767B910F CRC64;  
 Query Match 30.8%; Score 992.5; DB 1; Length 760;  
 Best Local Similarity 39.9%; Pred. No. 1.6e-75;  
 Matches 228; Conservative 91; Mismatches 210; Indels 43; Gaps 15;

OY 28 VPKRTVSGELAVVRFSGTQIDPLTLPLEPTGLVYGAREALFAFSE--ALEIQ 84  
 DB 39 MPKVKYHAGDGHRLSEFQOKGLRDPDTLLSDGNTLVYGAREYALANTIQGRLPK 96  
 OY 85 GAISFAPVPEKTECIQKGNKNOTECFNTIRFLQPNASHLYVGYARQPKTYNM-- 142  
 DB 99 NMIPVASEKRTKECAFKKNSMETQCFNFRVLYSYNAHLYVACGFAPSPACTFELAD 158  
 OY 143 -LFTLEHGEFEDEKCKCPDPKAGHAGLVVDLELSATLNNLTGEPITLRMGHHNS 201  
 DB 159 SLILPFLIDVMDKQSPFLFTSTQAVLYDGMLSGTNNPLGSEPLIMRTLGSHPV 218  
 OY 202 KTEYLAFWLN-EPHFVSAVVPESVGSFTGDDKRVYFFERAVESDCYAEQVAVAV 260  
 DB 219 KTDIFLRWLHADASFV--AAIPST-----QVYFFEFETSEDFEELISVAVQ 268  
 OY 261 CKGDMGAGARTLQKRTTFLKARLACAPNQLYFNOLQAMHTLQDTSMHTTFGYVQAO 320  
 DB 269 KNDVCGEKILQKMTTFLKALQCAOP-QLPFNIRHVVLLPASPVSRYAVFTSQ 327  
 OY 321 W--GDVYLSAICETQLEIORVEGPKYKEHAEAOQMDRTDPVSPRQSCINNHRH 378  
 DB 328 NQVGGTRSSAVCFSLTDIERVYKKEKELKETSMTTYTRGSEVPRQSC----- 379  
 OY 379 GYTSLSLPDNLNLFVKKRPLMEQYQPRMSRPLVKKGTNFTLVADRYTGIGATYTV 438  
 DB 380 ---SMGSSSDKALTFPKHDFLMDENH---VGTPLVKSQVETRYLAVESARGLDGSHVY 433  
 OY 439 LFTGTDGHLTKA-VSLGPRVHLIELOLF-DQEPKRSVLSQSKLLPAGSGLVQLP 496  
 DB 434 MYLGTSTGPHKRVVQDSSAVLYVEBIQLSPDEPVRNQLAQAQAVAFSGGIWVP 493  
 OY 497 VADCIKTRSCADCVLADPVCAMSVNTSRCAVAGHFGSLILIOHWTSDTSGICN----L 552  
 DB 494 RANCSYTESCYCVLADPVCAMSVNTSRCAVAGHFGSLILIOHWTSDTSGICN----L 552  
 OY 553 RGSKIQSGP-XPKNTIVVAGTDLVLPCHLSS 582  
 DB 553 RSPRROSPPOLIKELVLPVNSIELRCPHLSA 584

RESULT 14  
 SMA\_MOUSE  
 ID SMA\_MOUSE STANDARD: PRT: 772 AA.  
 AC 008665; 062180; 062215;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D)  
 GN SEMA3A OR SEMAD OR SEMD.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_Taxid=10090;  
 RX (1)  
 RC SEQUENCE FROM N.A.  
 RX STRAIN-NMRI: TISSUE-Embryo;  
 RX MEDLINE=95267431; PubMed=7748561;  
 RA Puschel A.W., Adams R.H., Betz H.;  
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family  
 and creates domains inhibitory for axonal extension.";  
 RL Neuron 14:941-948(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.



CC  
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC  
CC -1- DEVELOPMENTAL STAGE: At E-11, expression was restricted to the  
CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELENCEPHALIC  
CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RAHWE'S POUCH, AND  
CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED  
CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES  
CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS  
CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC  
CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN  
CC MITRAL CELLS, NEURONS OF THE ACCESSORY BUDD AND CEREBRAL CORTEX,  
CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND  
CC SPINAL MOTONEURONS.  
CC  
CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY  
CC THIRD OF THE PROTEIN.  
CC  
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC  
CC -1- SIMILARITY: Contains 1 sema domain.  
CC  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----

CC	EMBL, X95286; CAA64607.1; .	
DR	InterPro: IPR007110; IG-like.	
DR	InterPro: IPR003599; IG.	
DR	InterPro: IPR003006; IG-MHC.	
DR	InterPro: IPR003659; Plexin-like.	
DR	InterPro: IPR001627; Sema.	
DR	Pfam: PF00047; Ig; 1.	
DR	Pfam: PF01403; Sema; 1.	
DR	SMART; SM00409; IG; 1.	
DR	SMART; SM00423; PSI; 1.	
DR	SMART; SM00630; Sema; 1.	
DR	PROSITE; PS50835; IG-LIKE; 1.	
KM	Signal: Immunoglobulin domain; Multigene family; Neurogenesis;	
KM	Developmental protein; Glycoprotein.	
FT	SIGNAL	1 20
FT	CHAIN	21 772
FT	DOMAIN	240 538
FT	DOMAIN	577 665
FT	DOMAIN	728 770
FT	DISULFID	650 723
FT	CARBOHYD	53 53
FT	CARBOHYD	125 125
FT	CARBOHYD	591 591
FO	SEQUENCE	772 AA; 88608 MW; 240907812FP9P22 CRC64;

```

Query Match      29.1%  Score 937.5; DB 1; Length 772;
Best Local Similarity 38.6%  Pred. No. 6.8e-71;
Matches 208; Conservative 95; Mismatches 199; Indels 37; Gaps 15;

QY      10 LAARLMGLGIGAEYWM-----NLVPRKTVSSGELATVVRRESQNGI---ODFLVLTLEPT 62
      1  | | : : | | : | | | : | | : | | : | | : | | : | | : | | : | | : | |
Db      7 IACLEMGILTLARANYANGKNVPRILKSYKEMESNNVITFENLAASSSYHTFLDEER 66
      1  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      63 GLLYVGRREALFAFSMALELQGLAISWAPYEKTECIQKGNKNQTECFNIFRLQPYNA 122
      1  | | | | : : | | : : | | : : | | : : | | : : | | : : | | : : | |
Db      67 SRLVVGAKRDIHFSFNLVNIKDFOKIVPVSVSTRDECKWAGKDKILKECANIKVLKAYNQ 126
      1  | | | | : : | | : : | | : : | | : : | | : : | | : : | | : : | |

QY      123 SHLVVCGSTVROPFCSTVYVM-----LTFLEHEFEFDGSGKCYDDAKAGAGLNGDEL 176
      1  | | | | | | | | | | | | : : | | : | | | | | | | | | | | | | | | |
Db      127 THLYACTGTGAHPHICTYIEVGHHPEDNIEFKLODSHFENGKRSYDRLKLTASTLIDGEL 186
      1  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      177 YSATLNNPLQCTEPILIRNMGPHHSMKTE-YLAETFLNPEHFGASAVVESSGFTGDDKRV 235
      1  | | : : | | : | | | | | | | | | | | | | | | | | | | | | | |
Db      187 YSGADADFQWRDFAIFRTLLGHHPHIREQDHSRLNDPRPISALILIESDNP---EDDKV 243
      1  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      236 YFFPRRLAAYSDCAEQYVAVARVCCGDMGCAFTLQKKTFPLKALACAP---NMQL 292
      1  | | | | | | : : : : | | : : | | | | | | | | | | | | | | | | | |

```

```

Db      244 YFFFRFNALDGSRGKATHARIQIOLCKNPGGHRSLVNWKMTPTLAKRLILCSVPGNGIDT 303
QY      233 YVNOQLAMHTLDDTSMHNTFFGCVYQAQNGDMTLSACEQLEIQRVGPPKRETHEEA 352
Db      304 HFDLIDQVFLMKSCKDPKNPITYGVFTSSNIEFGSVACMTSMDSVDRVFGPPAHNDGPN 363
QY      353 QKMDRYTDPVPSPRPDQSCINNHRHRRGYTSSELPDNIILNFVKRHPRLMEQVGPBRMSRPL 412
Db      364 YQWVYFGVGRVPRPRGTCPSKTF--GGFSDTKDLPDDVITFARSHBAMVNPVPIINRPI 421
QY      413 LVKKGSTN--FTGLVADRYVGTGLGATVTVLFIFIGDGMILKAVSL--GPVH-----LIEE 463
Db      422 MKRTDVNTQFTVIYVDRVDAEDG-QYDAVFIGTVSTVLKVSVPRETM-HDLEEVLLER 479
QY      464 LQLFDQEP--MSLVLISQSKLLFAGSRQVLQVLPVADCIKY-RSCADCVIARPYCAM 519
Db      480 MTFV-REPPTISAMELSTRKQOOLYIGSTAGVQALPLRHRRDIYQKACACECLARDPYCAM 537

Search completed: August 7, 2003, 13:48:02
Job time : 28 secs

```

Search completed: August 7, 2003, 13:48:02  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:45:14 ; Search time 105 Seconds  
(without alignments)  
1464.756 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219  
Sequence: 1 MASHMAVWLLAARLWGIG.....PCHLSSNLAIPDSNPESSV 596

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP archaea:\*
- 2: SP bacteria:\*
- 3: SP fungi:\*
- 4: SP human:\*
- 5: SP invertebrate:\*
- 6: SP mammal:\*
- 7: SP mhc:\*
- 8: SP organelle:\*
- 9: SP phage:\*
- 10: SP plant:\*
- 11: SP rodent:\*
- 12: SP virus:\*
- 13: SP vertebrate:\*
- 14: SP unclassified:\*
- 15: SP virus:\*
- 16: SP bacteriophage:\*
- 17: SP archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	90.6	963	4	O9C0C4
2	1353.5	42.0	510	4	O9NXY2
3	1240	38.5	893	4	O9C0B8
4	1179.5	36.6	823	11	O8BIR6
5	1178.5	36.6	799	11	O8BUC1
6	1169.5	36.3	550	11	O8BIC3
7	1029.5	32.0	761	4	O8WUA9
8	913	28.4	774	13	O8JTW9
9	884.5	27.5	748	4	O8B7I1
10	878.5	27.3	777	11	O8B834
11	876.5	27.2	754	4	O8BDV7
12	872.5	27.1	777	11	O8BWF6
13	852.5	26.5	775	11	O9OXK3
14	851.5	26.4	782	4	O9NS98
15	818	25.4	756	13	O9OCU9
16	779	24.2	635	4	O96GX0

17	730.5	22.7	1202	4	O9P283	O9P283 homo sapien
18	729	22.6	998	4	O8NFX6	O8NFX6 homo sapien
19	729	22.6	1011	4	O8NFX3	O8NFX3 homo sapien
20	729	22.6	1017	4	O8NFX5	O8NFX5 homo sapien
21	729	22.6	1022	4	O9P249	O9P249 homo sapien
22	729	22.6	1073	4	O8NFX4	O8NFX4 homo sapien
23	727	22.6	595	11	O8BKG8	O8BKG8 mus musculu
24	726.5	22.6	935	4	O96JF8	O96JF8 homo sapien
25	726	22.6	844	11	O8BXU8	O8BXU8 mus musculu
26	706.5	21.9	963	11	O91X36	O91X36 mus musculu
27	693	21.5	1049	4	O9P2H9	O9P2H9 homo sapien
28	686	21.3	1030	4	O9H2E6	O9H2E6 homo sapien
29	683	21.2	697	5	O8MLP1	O8MLP1 drosophila
30	681	21.2	724	5	O9V707	O9V707 drosophila
31	672.5	20.9	605	11	O8BXZ7	O8BXZ7 mus musculu
32	672	20.9	587	11	O8BUT0	O8BUT0 mus musculu
33	646	20.1	1005	11	O9E071	O9E071 mus musculu
34	645.5	20.1	476	4	O8NFX7	O8NFX7 homo sapien
35	645	20.0	403	11	O8BK06	O8BK06 mus musculu
36	641.5	19.9	687	4	O9BXR8	O9BXR8 homo sapien
37	637	19.8	202	13	O8UYD5	O8UYD5 xenopus lae
38	635	19.7	770	5	O44253	O44253 drosophila
39	634	19.7	770	5	O9V3M4	O9V3M4 drosophila
40	633.5	19.7	616	5	O9V7P8	O9V7P8 drosophila
41	615	19.1	457	4	O9HBR1	O9HBR1 homo sapien
42	614	19.1	367	4	O9HAR9	O9HAR9 homo sapien
43	614	19.1	418	4	O96JF9	O96JF9 homo sapien
44	606.5	18.8	923	11	O8RAU3	O8RAU3 mus musculu
45	604.5	18.8	920	11	O8RAU4	O8RAU4 rattus norv

## ALIGNMENTS

RESULT 1

O9C0C4 PRELIMINARY; PRT; 963 AA.

AC O9C0C4;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein KIAA1739 (Fragment).

GN KIAA1739.

OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-21082932; PubMed-11214970;

RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XIX.

RT The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro.";

RL DNA Res. 7:347-355(2000).

DR EMBL; AB051526; BAB21830.1; -

DR Genew; HGNC:10731; SEMA4C.

DR InterPro; IPR003599; IG\_1.

DR InterPro; IPR007110; IG\_1like.

DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR002165; Plexin-repeat.

DR InterPro; IPR001627; SEMA.

DR Pfam; PF01437; PSI; 1.

DR Pfam; PF01403; SEMA; 1.

DR SMART; SM00409; IG; 1.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00630; SEMA; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.

KW Hypothetical protein.

FT NON\_TER

SO SEQUENCE

Query Match 90.6%; Score 2918; DB 4; Length 963;

QY	8	W-LLAARIIMGICIGAEVMMNLVPRKTVSS-----GELATVVRPRFSQICIDPLTLTLTE	60
Db	133	WHLIYVA--WGAGSSRGARLRAVEPOGSCPSAAMLTPAEIATVRRRSOTGIDOLFLLTLTE	190
QY	61	PTGLIYVGARALAFPSMEALIELGAIWEAPVEKTECIGKKNNOFFCFIFIPLOY	120
Db	191	PTGLIYVGARALAFPSMEALIELGAIWEAPVEKTECIGKKNNOFFCFIFIPLOY	250
QY	121	NASHLYVCGTAAFOPKCTYVNMALFTTLEHGEFEDKGRCPYDPRAKSHGLLVDELGYSAT	180
Db	251	NASHLYVCGTAAFOPKCTYVNMALFTTLEHGEFEDKGRCPYDPRAKSHGLLVDELGYSAT	310
QY	181	LNMLLGTFTPIILRNMGPHHSKMTETYLAWLNEPHVGSAYVPSVGSFTGGDDKYYFFFR	240
Db	311	LNMLLGTFTPIILRNMGPHHSKMTETYLAWLNEPHVGSAYVPSVGSFTGGDDKYYFFFR	370
QY	241	ERAEVSDCYAQQVAVARAVKCGKMGARTLORKWTFLKARLACAPRMOLYFNOLAM	300
Db	371	ERAEVSDCYAQQVAVARAVKCGKMGARTLORKWTFLKARLACAPRMOLYFNOLAM	430
QY	301	HTLDDTSVHNTTFEGVFOAQMGDMTLSAICEYOLEIQRVEBPYKEVHEEAOKMDRYTD	360
Db	431	HTLDDTSVHNTTFEGVFOAQMGDMTLSAICEYOLEIQRVEBPYKEVHEEAOKMDRYTD	490
QY	361	PVSPRPPOSCINNMRHRCGYTSSLELPNILNMFVKKHPLMEQVPRMSRPLLYVKKGTNF	420
Db	491	PVSPRPPOSCINNMRHRCGYTSSLELPNILNMFVKKHPLMEQVPRMSRPLLYVKKGTNF	550
QY	421	THLVADRYTGLDGTATYVLEFGTGGDWLLKAVSLGPMWHLIELOLFDQEPMSRSLVSQS	480
Db	551	THLVADRYTGLDGTATYVLEFGTGGDWLLKAVSLGPMWHLIELOLFDQEPMSRSLVSQS	610
QY	481	KLLFAGSRSLDLOPVADCKIKYRSCACQVTLARDPYCAMSVTSCVAVNGHFGSLTIOH	540
Db	611	KLLFAGSRSLDLOPVADCKIKYRSCACQVTLARDPYCAMSVTSCVAVNGHFGSLTIOH	670
QY	541	VMTSDTSGICMLRSGSKIQSGPXPKNITVAVAGTDLVLPCHLSGNLA-----LPDSN	590
Db	671	VMTSDTSGICMLRSGK-KVPRTPKNITVAVAGTDLVLPCHLSGNLAHAWTFCGRDLPALQ	729
QY	591	P 591	
Db	730	P 730	

Query Match	Best Local Similarity	42.0%	Score 1353.5	DB 4	Length 510
Matches 257	Conservative	2	Mismatches 8	Indels 11	Gaps 2
QY	324	MYLSAICCYOLEEIQRYVEGPGYKEYHEBAQKMDRTYDVPSPRGSCINNMRHRYSS	383		
DB	1	MYLSAICCYOLEEIQRYVEGPGYKEYHEBAQKMDRTYDVPSPRGSCINNMRHRYSS	60		
QY	384	LELPDNLINFPKKNHLMEEQYGRPSRPLVKKGTNPFNLVADRYTGIDGATTVYLFGT	443		
DB	61	LELPDNLINFPKKNHLMEEQYGRPSRPLVKKGTNPFNLVADRYTGIDGATTVYLFGT	120		
QY	444	GDGMLLKAVSGPWNHLEELQLEFQDEPMRSVLVSQSKLLFAGSRSLVOLPYADCIKY	503		
DB	121	GDGMLLKAVSGPWNHLEELQLEFQDEPMRSVLVSQSKLLFAGSRSLVOLPYADCIKY	180		
QY	504	RSCADCVLARDPYCAMSVMYTSRCVAVGSHFSSLLIQHWMTSDTSGICNLKSGKIQSGRP	563		
DB	181	RSCADCVLARDPYCAMSVMYTSRCVAVGSHFSSLLIQHWMTSDTSGICNLKSGK-KVRRTP	239		
QY	564	KNITVYAGTDVLVPCHLSSNLA-----LPDSNP 591			
DB	240	KNITVYAGTDVLVPCHLSSNLAHARMTGGRDLPAEOP 277			
RESULT 3					
Q9C0B8	PRELIMINARY:	PRT:	893 AA.		
ID	Q9C0B8.				
NC	09C0B8.				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)				
DE	Hypothetical protein KIAA1745 (Fragment).				
GN	KIAA1745.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE:21082932; PubMed:11214970;				
RA	Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;				
RT	"Prediction of the coding sequences of unidentified human genes. XIX.				
RT	The complete sequences of 100 new cDNA clones from brain which code				
RT	for large proteins in vitro."				
RL	DNA Res. 7:347-355(2000).				
DR	EMBL; AB051532; BAB31836.1; -				
DR	InterPro; IPR003659; Plexin-like.				
DR	InterPro; IPR002165; plexin_repeat.				
DR	InterPro; IPR001627; Sema.				
DR	Pfam; PF01437; PSI; 1.				
DR	Pfam; PF01403; Sema; 1.				
DR	SMART; SM00423; PSI; 1.				
DR	SMART; SM00630; Sema; 1.				
KW	Hypothetical protein.				
FT	NON_TER 1				
SO	SEQUENCE 893 AA; 98361 MW; D6C6C48DEE524F14 CRC64;				
QY	Query Match	42.0%	Score 1353.5	DB 4	Length 510
	Best Local Similarity	92.4%	Pred. No. 6.9e-113		
	Matches 257	Conservative	2	Mismatches 8	Indels 11
					Gaps 2
DB	25	WNLVPRKTVSSGELATVAVRPSQGIODELTLLTEPTGLLYVGAREALFAFSKALELQ	84		
DB	98	WALSPRISLPGSSERPEPLREACHISYV7ALLSBDGRITLYVAREALFALSSNLFLP	157		



```

0Y 85 GA-----ISEAPREKTEBCTIOGKNNOCTCFNTRFLODYNMASHVCGYAPORPCITY 140
0Y 158 GGEYQELLMGAAEKQKQCSFKKQDRBOCQNTIKTLPLBSHGLFTCGTAAPSPCITY 217
0Y 141 NMLFTLHGE-----FEDQKGCOPYDPAKGNAGLLVDGELYSATLNNFTLGPETILRN 195
0Y 218 NNNENFTLADDEKGNVLLEDKCGKCPEDPNTKSTALVADDELITYGVSSFOGNDPAISRQ 277
0Y 196 GPHHSKTEYLFWLNEPHFGSAAYBESVSTGGDDVYFEFFERBAVESOCYAOYVA 255
0Y 278 SLRPT-KTSSSLNNLODPAFWASATYPSBLSIGSGDDDIYFEFSESTGOEFEEFNATVS 336
0Y 256 RVARVCGDMGAGARTLORRWTFKLARLACSAPNMOLYENOLQAMHTLODT--SWHTTF 313
0Y 337 RIARICKGDEGERVYLOQRWTSFLKQOLCSRPDDPEFNVYLODVFTLSPSDQMDHTLE 396
0Y 314 PCVFAQW--GDMYLSAICEYOULEQRYPEBGRKRYHEAOKMWDYTPDPVSPRGSC 371
0Y 397 YCFPTSQMRGTTEGSAVCFTPKDQORVPSGLYKRYVNNRTOQWYTVHVPYPRGAC 456
0Y 372 NNMHRHGTSSLELFDNLINFEYKHPHLEDEOQPRMSRPLLYKKGTFNHYLADRYTGL 431
0Y 457 TNSARERKINSSLOLPDRVYLNKDFLKHDFLMDQGR--SRMLLLQPARQORAVAHVYPL 513
0Y 432 DQATYVYLITGQDGHLLKAVSLGPPVHLLIEQLDQ--EPMSLYLSQSKLLFPGSKS 490
0Y 514 H-HTYDLVFLGTGDXLHRAVSQGPVNHILIEQLTIFSSQOPVONLLDTHRGILLYAASH 572
0Y 491 QLVOLFVACICYYRSCADVLARPYCANSVYTRSCVAAGHNGSL---IDHWTSPT 546
0Y 573 GVVQVPMANCSLYRSCGCDLLARPYCANS--GSSCKHSLVQPOLATRWIODIGASA 630
0Y 547 SGICN-----LRSGKIOSGPXP-KNTIVVAGTDLVLPCHLSSNIA 585
0Y 631 KDLCSASSVYSPBFPVPTGKPCQOVQOPRYVYVWTATLACPLLSNIA 674

```

```

RESULT 4
OBIR6
ID OBIR6 PRELIMINARY; PRT: 823 AA.
AC OBIR6;
DT 01-MAR-2003 (TREMBlrel, 23, Created)
DT 01-MAR-2003 (TREMBlrel, 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel, 23, Last annotation update)
DE Semaphorin 4B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI-TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium.
RT The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK030331; BAC26906.1; -.
SQ SEQUENCE 823 AA; 91485 MW; E3B8C5E5CPALCD45 CRC64;

Query Match          36.6%; Score 1179.5; DB 11; Length 823;
Best Local Similarity 43.1%; Pred. No. 6,1e-97;
Matches 239; Conservative 92; Mismatches 215; Indels 35; Gaps 12

QY 7 VLLAARLNGLCIGAEVMMNVLPRKTVSSGSELATVVRRSQTGIODFLTLTLTEPTGLLY 66
   | | | | | | | | | | | | | | | | : : | | : : | | : : | |
Db 16 VILLLLLTLWTTTRR-----LGPRINVPCLCGEERLRKRREANISNYTALLLSQGKRLY 70

QY 67 VGAREALAFPSMALELOGA-----ISMENPVEKKECTICIOGKNNOETECNEFRLOPYNA 122
   ||||| | | | | | | | | | | | | | | | | : : | | : : | | : : | |
Db 71 VGAREALAFPLNSLSTLPGEYOEELIWSADARKQQCSRGKDPRDCONYTKILPLNS 130

```

[illegible]

```

RESULT 5
Q8BUCI
ID Q8BUCI PRELIMINARY; PRT: 799 AA.
AC Q8BUCI;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466651;
RA THE FANTOM Consortium,
RA "The Riken Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
RL RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK088653; BAC04080.1; -.
FT NON_CODING
          1
SC SEQUENCE      799 AA;    88814 MW;   415D3E687150A59A CRC64;

Query Match                36.6%; Score 1178.5; DB 11; Length 799;
Best Local Similarity     48.1%; Pred. No. 7.2e-97;
Matches 255; Conservative 74; Mismatches 182; Indels 19; Gaps 11,

OY        66 YVGAREALFAF-SMEALELOGAISWEAVEKKECTIOGSKNNOPCEFPFIIFLOPYNASH 124
           |||||...|| : : | . . . : ||::|| ::|| ||| : : -
DB         1 YVGAENEAFVANVALINSEKHIEEVTKVSDEDKSKCAEKSGKSQCETCLYNIRIVLDPLSRTS 60
           ||||||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY        125 LVVCSTVAEPKCPTVVNNLTFTLEHGEEDGDCKPCYPDPAGHGALLVDGLYSATLNFF 184
           ||||| |||| : : : : : ||||| |||| : : : ||||| |||| : : : ||||| |||| : : :

```

	Query Match	36.3%	Score 1169.5	DB 11	Length 550	
	Best Local Similarity	46.2%	Pred. No. 27e-96			
	Matches	239	Conservative	87	Mismatches 172	Indels 19 Gaps 8
OY		27	LVPKRTVSSGELATVVRPFSQTDIFLTILTEPTGLLYGCAREALPAFMSBALRLOGA	86		
		:           :				
Dd		31	LCPISTVPLSGEERIRLFEAEINISNYALLISOGKTLTVGARALEPALNSNLSPGG	90		
		:           :				
OY		87	---ISWEAPPEKTECIQGRNMOTECENFRFDLPYNASHLLVYGCTAFOPKCYVMN	142		
		:           :				
Dd		91	EYGEIIMSGADARKOOCSEFKCDPRDCONYIKILLPLMSSHLICGTAAFPCLAYIH	150		
		:           :				
OY		143	LTFPLEHSE-----FEDGKGCPYPDPAHGAGLVWDGELYSLTNLNFCOTEPTILIRNMKP	197		
		:           :				
Dd		151	ASFTLEDEAGNVIILEDGKRCFPDPNRKSTLVLDGELYITVSFGQNDPAISRQSS	210		
		:           :				
OY		198	HHSMTLEYLAFLWLNPHFVGSAIVESVGSFTGDDDKYFFERRRAVESDCAEDOVAVY	257		
		:           :           :				

	Query Match	32.0%	Score 1029.5	DB 4	Length 761
	Similarity	40.2%	Pred. No. 176-83		
	Best Local				
	Matches	237	Conservative	89	Mismatches 217
					Indels 47
					Gaps 15
QY	28	VPKRVSSGELATVYRRPSQTGIODFTLTLETPTGLLYVGAREALFAFSME--ALEIQ	84		
DB	39	MPRVYVYAGDERBRALSPFHQKGIADPDTLLSGDNTLYVGAREALIALDIODDPVRLK	98		
QY	85	GAISAEAVRRKEETEDIOGKKNNGPSCFPIRIQLQYNSHLVCGTAYAPQCTVYNM--	142		
DB	99	NMIPAPADNRKSCBAFNRKKSNEYCFNFIIRLVSYNTHLTYCGTFAPASPACTELEDD	158		
QY	143	-LTFLEHGEEEDGCKCPRYPAPAKHAGLVLGVSELYSATLNLFTGPRIILNMPRHNSM	201		
DB	159	SYLLPISIDKWEKSGQSPRPRAHKHTAVLDVGMLYGTMNMFLESEPTLMTLSQPYL	218		
QY	202	KTEYLATFL-NEPHVGSATVPESVGSFTGDDDKYVFFPRKRAVNSDCYAEQVAVARVY	260		
DB	219	KTNDFLRLMHHDAQSV--AAIPST-----QVVYFEEELASGDEFERLHTSVAVY	268		

```

OY      261  CKGDGSGARTLORRKTFTFLKARLACSAFNMQLQLOAMHTLQDTSMWNTTFEGVFOAO 320
Db      269  CKNDYGGKRLQKRTTKTLKQLDLCTQGG-QLPENVINHAUPLPADSPAPATITAVFTSQ 327
OY      321  W-GDMYLSAICEVQLEIQRVFEGRPYKEYHEBAQKMDRYDTPVPSPRGSCINMHRRH 378
Db      328  MGVGSTRSSAVCAFSLLDIERVEFKKYEKLNETSRMTWTYRGPETINPRGSC----- 379
OY      379  GYTSLELPLDNILNPKKHPHMEQGYGRWSRPLLYKGTNTHTLVADRVTGLDCAVTV 438
Db      380  ---SYGPPSSDALPTFMKHFPLMDEGV---VGTPLLYKGVETRTLAVETAQSLDHSILY 433
OY      439  LEFGDGMILKAVSLIG-PWVHLIEQLDF-DOEPMRSIVLSQSKLLEFAGRSQLYOLP 496
Db      434  MYLGTTTSLKRAVYSGDSASHLYBEIQLFPPEPVRNQLAPLPGAVFYVGGSGVMKRP 493
OY      497  VADCIKRYSCADCVLARDPYCAMSVNTSRCAVAGVGHFSGLLIQHWYTSIDTSGIC---NL 552
Db      494  RANCSYVESYDCVCLARDPHCAMDESRCCLLSAPNLNISMQDMERGNPEMACSGPMS 553
OY      553  RGSRTQSGP-APKNTTVYAGDVLVPC-HLSNLLA-----PDSNPEBS 595
Db      554  RSLRQSPSPQIIRKEVLAVPNSLLELPCPHLSALASYWSHGPAVPEASS 603

RESULT 8
OY      08JTW9  PRELIMINARY;  PRT;  774 AA.
Db      08JTW9
AC      08JTW9;
DT      01-OCT-2002 (TREMblrel. 22, Created)
DT      01-OCT-2002 (TREMblrel. 22, last sequence update)
DT      01-MAR-2003 (TREMblrel. 23, last annotation update)
DE      Xenopus 3A.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OC      Xenopodinae; Xenopus.
OX      NCBI_Taxid=8335;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Head;
RA      Tannahill D., Nielsen J., Regan A.G.;
RT      *Xenopus semaphorin 3A.*;
RL      Submitted (Apr-2001) to the EMBL/GenBank/DBD databases.
DR      EMBL: AY030051; AAK36166.1; -.
DR      InterPro: IPR003599; Ig.
DR      InterPro: IPR007110; Ig-like.
DR      InterPro: IPR003006; Ig_MHC.
DR      InterPro: IPR001005; Myb_DNA_binding.
DR      InterPro: IPR003659; plexin-like.
DR      InterPro: IPR001627; Sema.
DR      Pfam: PF00047; Ig; 1.
DR      Pfam: PF01403; Sema; 1.
DR      SMART: SM00409; Ig; 1.
DR      SMART: SM00423; PSI; 1.
DR      SMART: SM00630; Sema; 1.
DR      PROSITE: PSS0835; IG_LIKE; 1.
DR      PROSITE: PSS0037; MYB_1; 1.
SQ      SEQUENCE 774 AA; 89319 MW; 3127ED2ECCD5FA4 CRC64;

Query Match      28.4%; Score 913; DB 13; Length 774;
Best Local Similarity 39.3%; Pred. No. 5, 2e-73;
Matches 204; Conservative 88; Mismatches 195; Indels 32; Gaps 13

OY      26  NLVPKRYVSSGGLATVVRFRSQTQ---IDFTLTLTPTPTGLLYGARGALPAFMSALE 82
Db      27  NNVPFLRLSLYKRMVESNNMLITNGPNSSYITFTVLDERGRLYYGAKDHIFSNIANYK 86
OY      83  LOGAISWEAPVEKTECIQGRKNNQTECFNFTIRFLQPYNASHLYYCYATAFQPKCTYVM 142
Db      87  EFQKILMPYNTINIRDBCKWAGADYKECANFIKLVKAYNHTHLNACGTGAFFHPCVYIDV 146
OY      143  -----LFTTLEHGEFEDEGKCKCPDPAKGAAGLLVDGELYSATLNNLNGTPEITLRNK 195

```

[illegible]

```

Db 132 GTGAHPCTAEEVGHRAEPEVRLDPPGIEDEKSGSPDPRIHRAASVLYGELTSGVAA 191
183 NFIAGTEPILRNMGPHSMKTE-YLAFLMNEPHEVGSAYVPESVSGFTGDDOKVYEFRE 241
192 DLAMGDETFIRSLGQRPRLRTERPHDSKMLNEKFKYAVFPISEBNP---DDDKITFFERE 248
242 RAVE-SDCYAEQVAVARAVAKGDMGAGATTQKWTTFELKARLACSAAPNMQ--LYENQLO 298
249 TVEAEPALGRLSVSVGQICRNDVCGSRSLVKNKMTTFPLKARLYSVPEVEGDTHTDQLO 308
299 AHHITQDTSMTHTTFEGVQAOMGDMYLSAICEYOLEEORVPEGVKEYEHEBAQMDXK 358
309 DVEFLSSRDHRTPLLYAVFST-SIFQSGAVCVYSNDVRAFLGFAHKEGPMQWASX 367
359 TDVPSPRGSCINNMHRRHG-YTSSLEPLDNILNFKHPLEBEQVGRMSPLVYKRG 417
368 QGRVYPRGKPS---KTFGTFSSRTKDDPDVIOGARHNPMLKYNVLTGGRPLEYQV 424
418 TN--FTHLVADRVTLGATYTVLFTIGTGMLLANVSLGPMVH-----LIEELQLF-D 468
425 ANVTFTQIADRVAAADG-HYDVLFTGTDTVLKAVISPKGSRPSAEGILLLELHVFD 483
469 QEPMSLYLSQSKLLFAGSRSLQVLPADCIK-RSCADCVLARDPYCAN-SYNTSRC 526
484 SAAVTSKQISKRHOLYASRSAAVAQIALHRCAGRGVTECCCLADPYCANDGVACTRF 543
527 VAVGHPGSLLIQHVNTSDTSGICMLRSGKISGHPKNTVYAGTDLVPC 578
544 QPSAKR--RFRQDVANDPSTLCS--GDSRPLLEHNVFVEGSSATLBC 591

```

## RESULT 10

```

Q8BH34 PRELIMINARY; PRT; 777 AA.
AC 08BH34;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney, and Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK028900; BAC26185.1;
DR EMBL: AK052671; BAC35092.1;
SQ SEQUENCE 777 AA; 89548 MW; 5450DBD45D1DBAF CRC64;

```

Query Match 27.3%; Score 878.5; DB 11; Length 777;  
 Best Local Similarity 38.6%; Pred. No. 6,6e-70;  
 Matches 193; Conservative 88; Mismatches 190; Indels 29; Gaps 14;

```

QY 46 SQTGODDLTLTLPRTGLLYGAREALFAFSEMLELQ-GAISEAPYEKTECTQK 104
DB 64 SSEGI-DEVTLLDERGILLGARDHVELLSLVLNKFKKIYPAKEREVELCKLKG 122
QY 105 NNQTECFNIRFLOPYNASHLYVCTYAFOPKCTVNM-----LFTLEHGEFEDEGK 158
DB 123 DANACAFIRKLOPYNKTHTYVCTGAHPRLCGIIDDGANKKEELFKLDTHNLSGRK 182
QY 159 CPYDPAKAGLLVDELVSATLNNFLCTEPIILNNMG---PHNSKTEYLA-FWLNPEH 214
DB 183 CPFDQGFASVMTDEHLYSGTASDFLKDFTFTSLGLMDHHSIRDISHHNLNGAK 242

```

```

QY 215 FVGSAYVPESVSGFTGDDDKVYFEFFERAVESDCVAEQVAVARVYCKGDMGARTQK 274
DB 243 FICTFPIPD---TYNDDDKIYFEFFRESSQEGSTSDRSILSRVGRKNDVGGQRLINK 299
QY 275 WTFELKARLACSNP---NNQLYFNQLOAMHTLQDTSMHTTFEGVQAOMGDMYLSAICE 331
DB 300 WTFELKARLICSIPGSDAGDTHEDLDIYLLPTRDERNPVYGVFTTSSIRKGSAYCV 359
QY 332 YOLEIORFEGEYKYEHEBAQMDRYTDPVPSPRGSCINNMHRRHGHTSSLEPDLN 391
DB 360 YSMADIRAFNGYFAIKESDHRNVOYDGRIPFRGTCPSKTYDL-1KSTDFPDVY 418
QY 392 NFVKKHPLMEEGVPRMSRPLLYKGTNF--THLVADRVTLGATYTVLFTIGTGMWL 449
DB 419 SFTRHPVYKSYVPAAGAPTFKRINVDYRLQIYVDHYVAEDG-QYDVWFLGTDTGTVL 477
QY 450 KAVSLG--FW---VHILEELQLFQD-EPARSLVLSQSKLLFPGSSQVLPVADCIK 503
DB 478 KAVSISKERNNEEVLLEQVEKHPRTAILNMLSLKQOOLYVSGMDGLVQLSHRCDTY 537
QY 504 -RSCADCVLARDPYCANSVN 522
DB 538 GKACADCCCLADPYCAMDGN 557

```

## RESULT 11

```

Q8TDV7 PRELIMINARY; PRT; 754 AA.
AC 08TDV7;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin 3B.
GN SEMA 3B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
RN 11;
RP SEQUENCE FROM N.A.
RA Koyama N.;
RT "Semaphorin 3B (SEMA3B) cDNA."
RT Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB083186; BAB88870.1;
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig-MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR01627; Sema.
DR Pfam: PF00047; Ig. 1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig. 1.
DR SMART: SM00423; PSI. 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PSS0835; IG_LIKE. 1.
SQ SEQUENCE 754 AA; 83691 MW; E61FD08C04E9A68E CRC64;

```

Query Match 27.2%; Score 876.5; DB 4; Length 754;  
 Best Local Similarity 35.7%; Pred. No. 9,5e-70;  
 Matches 213; Conservative 93; Mismatches 248; Indels 43; Gaps 19;

```

QY 12 ARLKGLIGAEYWMNLVPRKTVSSGELAT--VYRFSQTGIDFTLTLTPTGLLYGA 69
DB 14 ALLMAYGLGSAA--DSPRLRLSFQELQAMHGLQFSLBRTCCYQALVDEDEGRFLVGA 71
QY 70 REALFAFSEMLELQ-GAISEAPYEKTECTQKNNQTECFNIRFLOPYNASHLYVC 128
DB 72 ENHVASLNDNISKRAKLLAMPAPVEMRECCMWAGADIGTECMNVYKLLHAANRHLIAC 131
QY 129 GTYAPQPKCTVYNM-----LFTLEHGEFEDEGKCPYDPAKAGHGLVDELVSATLN 182
DB 132 GTGAHPCTAEEVGHRAEPEVRLDPPGIEDEKSGSPDPRIHRAASVLYGELTSGVAA 191

```

```

QY 183 NFICTEPIILRNKGRHHSKTE-YLAFWLNBPYVGSAYVPSVSTGDDDKYFFFR 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 DLNGRDTIERSLQRPRLTEPPDSRLMEPKRVKWIPESEN- -DDDKTYFFFR 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 RAVE-SPCYAQQVAVARVAKGDMGARTLQRTKTEFLKARLACAPNMQ--LYFNOL- 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 TAVEAALRLSLYSRQGLCRNDYVGSORSIYKMTFTFLKARLACVSPVGECDTHFQDLR 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 ----QAMHTLQDSWNTTFFGVFOAQMGDMYLSAICEYOLEEIQREYKPEYHEAQ 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 PPAEDVFLLSNRDRTPLLYAVSTSSIFQGSACVCSYMNDRAPFLGFAKBEGRMH 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 KMDRTIDVPSPRPGSCINNMHRRHG-YTSSLELDPNINLFVKHPLMEQVGRMSRPL 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 QWYSQGRVYPRPGMCPSP--KTFGFSSTKDEPDVIOFARNHPLMNSVLPFGRL 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 LYKGTN--PTHVADVTGLDGTATVTLFGTGDGLLKAVALSGPVH-----LLEL 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 426 FLOVGANYTTQIADVAAADG-HYDVLGTGVGLVAVKISVPKGRSRSABGLLEEL 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 QLF-DEPMRSLVLSOSKLLFAGRSQOLVOLPYADCIKY-RSCADCVLARDPYCAM-SV 521
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 485 HVEFDSAAYTSMQISSRHLQIVASRSVAVQIALHRCANHRVCTECCIADRPICANDGV 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 522 NTSRCVAVGHFSGLLIQHWTSPTSGICNLGRSGKISGPRKNTVYAGTDLVLP 578
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 545 ACYHFGQSAKR--RFRQDVANGDPSTICS--GDSRPALLEHNVFVGESSAFLEEC 597
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 12

```

Q8BMF6 PRELIMINARY; PRT; 777 AA.
ID Q8BMF6
AC Q8BMF6;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
   60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK031704; BAC27522.1;
SQ SEQUENCE 777 AA; 89562 MW; B7204D8228B89CE CRC64;

```

```

Query Match 27.18; Score 872.5; DB 11; Length 777;
Best Local Similarity 38.48; Pred. No. 2.3e-69;
Matches 192; Conservative 88; Mismatches 191; Indels 29; Gaps 14;

```

```

QY 46 SONGIOEPIITLPEPTGLLYVGAREALFAFMSMALEIQ-CAISMARPVEKTECICOK 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 SSEL-POYLLDEERGLILGAKDHVFLSLVDLNNKFKIYPAKKEVVELCKLAGK 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 105 NNQTECFEIRFLOPYNASHLYVCGTAFQPKCTYVNM-----LFTLIEGEFEDGK 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 DANECANFIRIVLOPYNKTHVYCGTGFHPLCGYIDLGANKRELIFLIDHNLNESGLK 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 159 CPYPAKAGGLVLDGELYSATLNNFLGTEPIILRNK- -PHSMKTEYLA-TWLNPH 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 CPFPDQPFASVMDHLISGTASDFGKTAFTFRLSLGMDHHSIRTDISEHWMLNCAK 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 FVGSAYVPSVSTGDDDKYVFFFERAVESDCYAEQVAVARVAVKCGMDGARTLQK 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 FIGTFPIPD- -TYNPDCKIYFFFRSSQSGSTSDKSIISRGVACANDVGGRSLNK 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 275 WTFELKARLACAP- -NKOLYFNQLOAMHTLQDSWNTTFFGVFOAQMGDMYLSAICE 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 300 WTFELKARLACISIPGSDADTHEDELIDYILPRDRKRPVYGVGFTTSSIFKSAVY 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 YOLEEIQREYKPEYHEAQKMDRTIDVPSPRPGSCINNMHRRRGYSSLEPNIL 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 YSMADIRAVFNGPFAHRESADHRVWQYDGRIPRPETGCSKYDPL-ISTREPPDVI 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 392 NFVAKHPLMEQVGRMSRPLLYKGTN--PTHVADVTGLDGTATVTLFGTGDGLL 449
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 SFIRHHPYKSVYPVAGAPTEFKRIINDYRLQTVYDHYVAEDG-QTDVFLGTIDITVL 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 450 KAVSLG--PW--VHLIEQLQFDQ-EPMSLVLSOSKLLFAGRSQOLVOLPYADCIKY 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 478 KVASISKEKMMERVEVLEELQVFKHPAILMELSLNQOOLYVSGMGLVQLSLHRDQY 537
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 504 -RSCADCVLARDPYCAMSVN 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 538 GKACADCVLARDPYCAMDGN 557
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 13

```

Q90X23 PRELIMINARY; PRT; 775 AA.
ID Q90X23
AC Q90X23;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE Semaphorin M-Semak.
GN SEMA3E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57/Black 6;
RA Miyazaki N., Furuyama T., Inagaki S.;
RT "A novel semaphorin, M-Semak which inhibits neural outgrowth from
   sensory neurons."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF034744; AAD01996.1; -.
DR MGD: MGI:1340034; Semak3e.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00423; PST; 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PS50835; Ig-LIKE; 1.
SQ SEQUENCE 775 AA; 89543 MW; 221E76F40498D4 CRC64;

```

```

Query Match 26.58; Score 852.5; DB 11; Length 775;
Best Local Similarity 36.68; Pred. No. 1.4e-67;
Matches 210; Conservative 86; Mismatches 216; Indels 61; Gaps 22;

```

```

QY 1 MAPHMAVWLAAARLGLIGAEVNM-----NLVPRKTVSGSELATVVRFR---SQTGIQ 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MAP--AGHILLLMDGHL--ELWTPGHSANPSYRLRLSHKELLELRNRSIPSPGLFL 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 52 DFLITLPEPTGLLYVGAREALFAFMSMAL-ELQALISWAPVEKTECICOKNNOTEC 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 DLHMLDEYQERFLFVGGRDLVYSLNLERVSDGYREIYWPSTAVKVECCINRGR-DANEC 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 ENFIRFLOPYNASHLYVCGTAFQPKCTYVNM-----LFTLIEGEFEDGKCPYDPA 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 ANYIRVLAHNYKTHLILCATGAFDPHCAFIVGHHSEPLFHLHSHSERGRGRCPPDPN 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 KGHAGLLVLDGELYSATLNNFLGTEPIILRNKGRPHHSKTEY- -LAFWLNBPYVGSAY 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 176 SFEVSTLGNELFACLYSDYWGGRDSAIRFSMGLCHITREHDERL---LKEPKFVGSTW 232
QY 221 VPEVSGETGDDDDKDYFFERERAVESDCYAEQVAVARVCCKGDMGARTLOKRTTFLK 280
Db 223 IPDNDR---DNNKMYFFETTERALAEANNNAHTIYRVGLICVNDMGGRILVNNKSTFLK 289
QY 261 ARLACSAAPW---QLYFNOLQAMHTLQDTSWHTNTFFGVFOAQWDMYLSAICEYOLEBT 337
Db 290 ARLVCSVGMGMDIDYFDELEDFVFLPTRDPKNPVIFGLFNTTSTNIFRGHACVYHMSI 349
QY 338 QVFECPY-----KEYHEAOKMDRYTDPVSPRPGSCINNMHRRHG---YTSLELPONI 390
Db 350 REAFNGPFAHKEGPRYH---WSLYEGKVPYPRPGSCAS---KNGGKVTGTQYPPDA 401
QY 391 LNFVAKHPLMEQVPRMSRPLLVKKG---TNEFHLVADRYTGLDQATYVLEFICTGDL 448
Db 402 IRFAMHPLMYQPIKPYKHKPILVKTGKYNLRQLAVRVEADG-QYDVLFIGDTGLV 460
QY 449 LKAVSL---GPWH-LIEELQLF-DQEPMSLVLSOSKLLFAGRSQVQLPVADCI 501
Db 461 LKVTIYNOEREMMEVILEELQIFKDPAPITSEISSKROOLYIGSASAVAVQVRFHCD 520
QY 502 KYRS-CADCVLARDPYCAM-SVNTSRCAVAVGSH 532
Db 521 MYSACADCCCLARDPYCAMDGTSCSRYPPTGSH 553

```

## RESULT 14

Q9NS98 PRELIMINARY: PRT: 782 AA.

```

AC Q9NS98 Q9H7Q3; (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin sem2 (Flj00014 protein).
GN SEM2 OR FLJ00014.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
RA Salto T.;
RT "Human semaphorin."
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 58-782 FROM N.A.
RC TISSUE=Spleen.
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB029496; BAA98132.1; -
DR EMBL; AK024425; BAB15715.1; -
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003006; IG-MHC.
DR InterPro: IPR003659; Plexin-1like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;

```

Query Match 26.5%; Score 851.5; DB 4; Length 782;  
 Best Local Similarity 36.6%; Pred. No. 1.8e-67;  
 Matches 204; Conservative 89; Mismatches 210; Indels 55; Gaps 16;

```

QY 1 MAPH-WAV-WLLAARLKLGLICAEVWMLVPKRVSSGELATVVRPSQTGIQDFLTTL- 57
Db 1 MAPSAMALCWLGLLGLLSSGSSGSPGSPVRLRLSLRDLSSANSAIFLPGQSLNLQA 60
QY 58 --LTPETGLLYGABEALEAFSMLEALLEGALISWEAPV-----KTEIOGKNQ 107
Db 61 MYLDEYRDLRLFLGGLDALYSRLDQ-----AMPDPREVLMPPQPGQRECVRRGRDPL 113
QY 108 TECPNFIRELPYRNASHLVGCTGYAFQPKCTYVNM-----LFTLEHGEFEDGKRCPD 162
Db 114 TECANFVRVLPQHNHTHLLAGCTGAFQPCALITVGHREHVLHLEPGSVESGRRCRPE 173
QY 163 PAKGAGLLVDCELYSATLNNFLGTEPITLRMGGRHSHKTEYLAFLWLEHPVGSAYVP 222
Db 174 PSRPASTFIDELLYTGLADFLGREAMIFRSGGPRPALRSDQSLHDPREFVAAR 233
QY 223 EIVSGFTGDDDKVYFFERERAVESDCYAEQV-VAVAVARVCCKGDMGARTLOKRTTFLK 281
Db 234 ENSDO---DNDKVTFFSETVPSPDGSHVTVSRKVCVNDAGQRLVNNKSTFLKA 290
QY 282 RLACSAF---NMQLYFNOLQAMHTLQDTSWHTNTFFGVFOAQWDMYLSAICEYOLEEIO 338
Db 291 RLVCVSPGPGAEETHEPQLEDFVFLMPKAGSLLEYALESTVSAVQGPVAVYHADIW 350
QY 339 RVEEPKYETHEAOKMDRYTDPVSPRPGSCINNMHRRHG---YTSLELPONILNPKK 396
Db 351 EVFNGPFAHRDGPQHQWGPYGGKVPFPRGVCPSKTAQPPRPFESTKDPDEVILQFANA 410
QY 397 HPLMEQVPRMSRPLLVKKGNTFHL-----VADRYTGLDQATYVLEFICTGDL 449
Db 411 HPLMFMPVPRHGRVLRK-----THLAQQLQIYVDRVREADG-TYDVIFLSTGSGVL 464
QY 450 KAVSL-----GPWH-LIEELQLF-DQEPMSLVLSOSKLLFAGRSQVQLPVADCI 502
Db 465 KYIALQAGSAPREVEVILEELQVFRVPRPTEMEISYKQMLYGSRLGVAQLRLHQERT 524
QY 503 Y-RSCADCVLARDPYCAM 519
Db 525 YGTACAECCLARDPYCAM 542

```

## RESULT 15

Q8QG09 PRELIMINARY: PRT: 756 AA.

```

AC Q8QG09
DT 01-JUN-2002 (Tremblrel. 21, Created)
DE 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Semaphorin 3F.
GN SEMA3F.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe Y., Nakamura H.;
RT "Axon guidance of the trochlear nerve by Sema 3F along mid-hindbrain
RT boundary."
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB072930; BAB88691.1; -
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003598; IG-C2.
DR InterPro: IPR003006; IG-MHC.
DR InterPro: IPR003659; Plexin-1like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.

```

KW Immunoglobulin domain.  
SQ SEQUENCE 756 AA; 85196 MW; C36754C02541ED88 CRC64;

Query Match 25.4%; Score 818; DB 13; Length 756;

Best Local Similarity 35.5%; Pred. No. 1.7e-64;

Matches 195; Conservative 86; Mismatches 217; Indels 52; Gaps 16;

```
QY 6 AVWLLAARLMGIGIAEVMWNL-----VPRKTVSSGEL-ATVVRRESQ--TGIDOF 53
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4 ATVLLMNTLLTIG-----WRAHGKDGVPPTPRVOLSFELKATGTAHFNFLNSDY 57
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 54 LTLTLTPTGLLYGAREALFAFSEALELOG-AISWEAPYEKTECIQKKNQTECFN 112
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 RILKEDHDHMYGSKDYLSLDLHINREPLIHPASQORIEECISGKNSNGECGN 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 113 FIFRLOPYNASHLVVCGTYAFAPKCTVYNM-----LFTLEHGFEFGKGCYPDPAGK 166
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 118 FIFRLOPYNASHLVVCGTYAFAPKCTVYNM-----LFTLEHGFEFGKGCYPDPAGK 177
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 HAGLLVDGELYSATLNNFLGTEPIILRNMGPHHSMT-ETLAFWLNEPHFVGSAYVESV 225
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 178 TVSALINBELYAGYIDFMGTDAIFRTMGKOTAMRTDOYNSRWLNDPAFVRAQLPD- 235
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 226 GSFTGDDDKYFFFRRAVESDCYAEQYVARVARVCKGDMGARTLQKRTTFELKARLAC 285
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 236 -SSERNDDKLYFFFRKSADAP-LSPGYSRIGRICLNDGHCCLVKNKMTFLKARLVC 293
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 286 SAP---NMQLYFNOLQAMHTLQDTSMNHTFFGVFOAOWGDMYLSAICEYOLEEIOVFEE 342
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 294 SVPGPDGIETHEDELQDYFIQOTDITKNPVIYAFVSAGSVFKSAYVCYSMADIRNVEFN 353
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 343 GPYKEYEHEAOKMDRYTDPVSPRPGSCINNMHRHGYTSSL---ELPDNILNFVKRHP 398
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 354 RPFHKKEGPNYQMMPTGKMPYRPGCPG-----TFTPSMKSTKIDYDEVINPMRSHP 408
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 399 LMEQVGPWRMRRLLYKGTN--FTHLVADRYTGLDGATYVLFITGDDGHLKAVSLGP 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 409 LMYHAVYPAHRQPLVVRTNVRRTTIAVDVDAAG-RYEVLFGTDRTGTVQKVIYLP 467
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 457 WVHLIEELQLEFDOE-----PMRSILVLSQSKLLFAGRSQLOVLPVADCIKY-RSCADC 509
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 468 DDMETEELMLEIEIVFAPAPIRKMTISSKQQLYSSAVGYTHLALHRCDYGEACADC 527
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 510 VLARDPYCAW 519
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 528 CIARDPYCAW 537
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 7, 2003, 13:49:56  
Job time : 109 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 7, 2003, 13:46:05 ; Search time 30 Seconds

(Without alignments)  
840,576 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219

Sequence: 1 MAHWAVWILLARLMLGIG.....PCHLSNIALPDSNPRESSV 596

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1220.5	37.9	862	4	US-08-556-422A-2
2	1149	35.7	607	4	US-08-556-422A-4
3	932.5	29.0	655	4	US-08-556-422A-3
4	932.5	29.0	771	1	US-08-121-713D-54
5	932.5	29.0	771	1	US-08-835-268-54
6	932.5	29.0	771	2	US-09-060-692-54
7	932.5	29.0	771	3	US-08-833-391-54
8	932.5	29.0	771	4	US-09-060-610-54
9	932.5	29.0	771	5	PCT-US94-10151A-54
10	890.5	27.7	425	4	US-08-556-422A-7
11	852.5	26.5	775	4	US-09-308-179B-1
12	727.5	22.6	930	4	US-09-254-594-6
13	719.5	22.4	295	4	US-08-556-422A-6
14	704.5	21.9	730	1	US-08-121-713D-58
15	704.5	21.9	730	1	US-08-835-268-58
16	704.5	21.9	730	2	US-09-060-692-58
17	704.5	21.9	730	3	US-08-833-391-58
18	704.5	21.9	730	4	US-09-060-610-58
19	704.5	21.9	730	5	PCT-US94-10151A-58
20	685	21.3	724	1	US-08-121-713D-62
21	685	21.3	724	1	US-08-835-268-62
22	685	21.3	724	2	US-09-060-692-62
23	685	21.3	724	3	US-08-833-391-62
24	685	21.3	724	4	US-09-060-610-62
25	685	21.3	724	5	PCT-US94-10151A-62
26	680	21.1	929	4	US-09-254-594-3
27	658	20.4	650	1	US-08-121-713D-60

28	658	20.4	650	1	US-08-835-268-60	Sequence 60, Appl
29	658	20.4	650	2	US-09-060-692-60	Sequence 60, Appl
30	658	20.4	650	3	US-08-833-391-60	Sequence 60, Appl
31	658	20.4	650	4	US-09-060-610-60	Sequence 60, Appl
32	658	20.4	650	5	PCT-US94-10151A-60	Sequence 60, Appl
33	654.5	20.3	887	4	US-09-077-940A-2	Sequence 2, Appl1
34	646	20.1	888	4	US-09-077-940A-4	Sequence 4, Appl1
35	596.5	18.5	712	1	US-08-121-713D-64	Sequence 64, Appl1
36	596.5	18.5	712	1	US-08-835-268-64	Sequence 64, Appl1
37	596.5	18.5	712	2	US-09-060-692-64	Sequence 64, Appl1
38	596.5	18.5	712	3	US-08-833-391-64	Sequence 64, Appl1
39	596.5	18.5	712	4	US-09-060-610-64	Sequence 64, Appl1
40	596.5	18.5	712	5	PCT-US94-10151A-64	Sequence 64, Appl1
41	562	17.5	477	1	US-08-136-922-2	Sequence 2, Appl1
42	471.5	14.6	634	3	US-09-041-236-2	Sequence 2, Appl1
43	471.5	14.6	634	4	US-09-771-467C-2	Sequence 2, Appl1
44	471.5	14.6	666	3	US-09-240-410-2	Sequence 2, Appl1
45	406.5	12.6	606	3	US-09-041-236-4	Sequence 4, Appl1

## ALIGNMENTS

RESULT 1  
US-08-556-422A-2  
Sequence 2, Application US/08556422A

Patent No. 6576754

GENERAL INFORMATION:

APPLICANT: HALL, Kathryn T.

APPLICANT: FREEMAN, Gordon J.

APPLICANT: SCHULTZE, Joachim L.

APPLICANT: BOUSIOTIS, Vassiliaki

APPLICANT: NADLER, Lee M.

TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES

FILE REFERENCE: DRN-005CR2

CURRENT APPLICATION NUMBER: US/08/556,422A

NUMBER OF SEQ ID NOS: 7

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 2

LENGTH: 862

TYPE: PRT

ORGANISM: Homo sapiens

US-08-556-422A-2

Query Match 37.9%; Score 1220.5; DB 4; Length 862;  
Best Local Similarity 46.0%; Pred. No. 5.4e-126;  
Matches 264; Conservative 89; Mismatches 188; Indels 33; Gaps 14;

QY	28	VPRKTVSSGELATVVRREFSGTIGIDFLFTLEPTGLYVGAREALFAF-SWEALELQGA	86
DB	27	IPRTWREHVLV--QHEPDITVYSALLSEDDITIXIGAREVFAVNALNISEKQHE	84
QY	87	ISWEAPVEKTECIQKGNNDTECFNTRFLQPYNASHLYCGTAFQPKCTYVNMJFT	146
DB	85	YVMKVEDSKKAKCAKSKQTECLNTRVLOPSATSLYCGTNAFPACDHLNLTGFK	144
QY	147	LEHGFEDGKRCPRDPKAGHGLYDGEISATINNLGTEPITLRKMGPHSKTEYL	206
DB	145	F-LGKNEDGKGCPRDPKASHYSTVAVDGLSGTISYNTLSRISN-SSHSPLRTEYA	202
QY	207	AFMLNEPFGVAVYVESVGSPTGDDKYVFFFRAYESDCAVQAVARAVKGMG	266
DB	203	IPMLNEPFGVAVYVESVGSPTGDDKYVFFFRAYESDCAVQAVARAVKGMG	262
QY	267	GARTLQRTWTFELKARLACAPAMQLYPNOLQAMHTLQDTSNHTTFFGVQAGMDYL	326
DB	263	GLRTLQRTWTFELKARLACAPAMQLYPNOLQAMHTLQDTSNHTTFFGVQAGMDYL	322
QY	327	SAICGQLELQORV--DEPYKE--YHEDAQMDKITDPVSPRGSCINNNHRHGTS	382
DB	323	SAVCAYNLSTAELVESHKQYKMTSTVEQSHTKWVYNGPVPKPRGACIDSEARAANYTS	382

```

OY      383 SLEEDNLTNLFVKKRPPBLBEOGVGPMKSRRLTKKGFNFHLYADRYTGADGAYVYFLPG 44.2
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      383 SLNLPDKTLQVYKDHPLMDOSTPTPLDNRRRLTKKDYNTQIYVDRLQALDGSYTDVMEVVS 44.2
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      44.3 TGDGMLLRAVSLGPPVHLEELQLF--DQPMKSLVLSQSK--KLTFAGSRSQLVQPLPVD 49.9
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      44.3 TDRGALHRAISLEHAHVHIIIEETQLRDPDEFEPVOTLLSSKKGRNFVYAGNSGVYQAPLAF 50.2
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      500 CTKRSCADCVLARPFCAMSVNYSRCVAV-----GGHFGSLLIQHVMTSDTSGICN 55.1
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      503 CGKHQTCDCCVLARPFCAMSPPTATCVALLHQTESPSRG-----LIDF--MSGDAS--VCP 55.4
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      55.2 LRGSRIQSGPPKNITVAVGTDVLVPCHLSSILA 58.5
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      55.5 DK-----SKGSYRQHFFKHGCT--AEKCSQSKSILA 58.3
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
US-08-556-422A-4
: Sequence 4, Application US/08556422A
: Patent NO. 6576754
: GENERAL INFORMATION:
: APPLICANT: HALL, Kathryn T.
: APPLICANT: FREEMAN, Gordon J.
: APPLICANT: SCHULTZE, Joachim L.
: APPLICANT: BOUSSIOTIS, Vassiliki
: APPLICANT: NADLER, Lee M.
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
: FILE REFERENCE: DEN-005CPA2
: CURRENT APPLICATION NUMBER: US/08/556,422A
: CURRENT FILING DATE: 1995-11-09
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 607
: TYPE: PRT
: ORGANISM: Mus musculus
US-08-556-422A-4

```

Query Match	35.78;	Score 1149;	DB 4;	Length 607;
Best Local Similarity	43.68;	Pred. No. 2.6e-118;		
Matches 247;	Conservative 89;	Mismatches 201;	Indels 30;	Gaps 11;

[illegible]

Db 419 LKRAVTLSSRVHTEELQFFPGGPVQVNLILDSHGCLLLVASSHSQVQVVPANCSLPTC 478

Qy 507 ADCVLNADPYCAMSVTSCRCVAVGGHPSGLLQIHWTSDDTSG-----ICNLRSK-----IO 558

Db 479 GDDLNRDPRYCAMTSGACRLASL-----YQPDLASRPTQDIEGAVKELCKSSSTKAFIV 555

Qy 559 SGPRKNTIVYVACTDLVLPCHLSSNLA 585

Db 536 PGRPCQVQIQPTVNTLACPLISNLA 562

```

: RESULT 3
: US-08-556-422A-3
: Sequence 3, Application US/08556422A
: Patent No. 6576754
: GENERAL INFORMATION:
: APPLICANT: HALL, Kathryn T.
: APPLICANT: FREEMAN, Gordon J.
: APPLICANT: SCHULTZE, Joachim L.
: APPLICANT: BOUSSETIOTIS, Vassiliki
: APPLICANT: NADLER, Lee M.
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
: FILE REFERENCE: DEN-005CPA2
: CURRENT APPLICATION NUMBER: US/08/556,422A
: CURRENT FILING DATE: 1995-11-09
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 655
: type: prt
: ORGANISM: Homo sapiens
: US-08-556-422A-3

```

Query Match	29.0%;	Score 932.5;	DB 4;	Length 655;
Best Local Similarity	38.4%;	Pred. No. 3.5e-94;		
Matches 208;	Conservative 94;	Mismatches 203;	Indels 37;	Gaps 15;

```

OY      8 WL--LAARLNGJGIGEVMM-----MLVPRKTVSSELAIVRRSPQIG---ODFLFLTL 58
Db      3 MLIRYICLFGVLLTRRANYQGNKNVPRKLSTYEMLESNNVITFPGNLSSSYHFFLL 62
OY      59 TEPTGLLYVGAREALFAFSMEALDELQALISWEAPEYKTECIGKKNQNECFNFIHQ 118
Db      63 DEERSRLHYVGAKKHIFSPDLVNIKDFOKIWPVSVSTRDECKKMGKDKILECACFIVLK 122
OY      119 PYASHLHYVGCTAFQPKCTYYNM-----LTFTLHEHFEEDGKQDYPAPAGHGILV 172
Db      123 AYQTHLHYACGICAFHPICTTYEIGHNEBDNIFLENSHFNFGCKSPYDPKLLTSLTL 182
OY      173 DGLYSATLNNFLGTEPILLRNMGRPHHSNKT-YLAEFLNPHFVGSAYVPEVSGTGD 231
Db      183 DGLYSGTAADEFGRGDFALFRTLGHHNHRIRTEQHSRLANPKFISHLISESDNP---E 239
OY      232 DDGLYFFPRPRVAESQAEQVYAAVAVVCGGDMGCAATLQKTKTFLKARLCSAP--- 288
Db      240 DDGLYFFPRPRVAESQAEQVYAAVAVVCGGDMGCAATLQKTKTFLKARLCSAPGN 299
OY      289 NMQLYFNQLAMHTLQDTSWMTTEFGVFOAMQDMYLSAICEYOLEIORVPEGRKEY 348
Db      300 GIDTHDELQDVFLEKFKDKPRVYVYGVFTTSSNFKCSAVCMSMDVRVFGLGPAHR 359
OY      349 HEBAQKMDYTRDVPSPRPGSCINNMHNRHCVTSLELDNLLFVKKHPLMEQVCPRM 408
Db      360 DGNVQWVYQGRVPRPRGTCPSKTF---GGDSTKDLDPDYITFASSHPRMIPVPRM 417
OY      409 SRDLVKKGTN--FTHLVADRYTGADGATYVYLFITGIDWKLAAVSL--GPWVH----L 460
Db      418 NRIIVIKTDVNOQETQIYVDRVDAEDG--QDYWMFIGTDVGVFLAVVSIPIRETHYLEVL 476
OY      461 IEELOLEFDEP--MSVLVSQSKLLFLFSGRSQVLQPLVADICITY-RSCADCVLARDPYC 517
Db      477 LEEHNVF--HEPRALISMETLSTQOQVLYGSTAGVAGOLPLHACDYLIGACAECCILARPYC 535

```

OY 518 AM 519  
Db 536 AM 537

## RESULT 4

US-08-121-713D-54  
; Sequence 54, Application US/08121713D  
; Patent No. 5639856  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,713D  
; FILING DATE: 13-SEP-1993  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)343-4341  
; TELEFAX: (415)343-4342  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 771 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-121-713D-54

Query Match 29.0%; Score 932.5; DB 1: Length 771;  
Best Local Similarity 38.4%; Pred. No. 4.6e-94;  
Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;

OY 8 WL--LAARLWGLGIGAEYWM---NLVPRKTVSSGELATVVRPSQTGI---QDFLTTL 58  
Db 3 WLFRIVCLFVGMVLLTARANYONGKNVPRKLSYKEMESNNVTFNFNLAASSSTHTFL 62  
OY 59 TEPGLLYVGARALAFMSMELELOGAISWEAVEVEKTEIOGKNNQTCFNPFRILQ 118  
Db 63 DEESRLLYVGAKDIIFSPDLVINIKDFOKIVPVSYTRDECKWAGKDKLECANFIVLK 122  
OY 119 PYNASHLYVCGTYVAFOPKCTIVNM-----LFTLEHGEFEDGKCKPYDPAKHAGLLV 172  
Db 123 AYNGTHLYAGSTGAFHPICTYIEIGHPEDNIFKLENSHFNGSKSPYDPRKILTAASLI 182  
OY 173 DGLIYSATLNFLETEPIILANMGRPHSMKTE-YLATVLANPHVGSAYVESVGSFTGD 231  
Db 183 DGLIYSATLNFLETEPIILANMGRPHSMKTE-YLATVLANPHVGSAYVESVGSFTGD 239  
OY 232 DDKYVFFPRERAVSDCYAEQVAVARVAVCKGDMGARTLQKRTTFPKARLACAP---E 288  
Db 240 DDKYVFFPRERAVSDCYAEQVAVARVAVCKGDMGARTLQKRTTFPKARLACAP---E 299

OY 289 NWLYENOLAMHTIQTSMNTFFGVGQAQMDMTLSAICEYQLEEIQRVEGPKKEY 348  
Db 300 GIDTFHDELQDVFLNFKDPKPNPVYGVFTTSSNIFKGSAYCVMSMSDVRVFLGPVHR 359  
OY 349 HEEAKMDRYTDPVPSPPRGSCINNRHRCYTSLSLEPDIILNFVYKHKPLMEQVQPRW 408  
Db 360 DGPYQWVYQGRVYPRPGTGPSTKTF--GGFSDTKDLDPDVITFARSHPMYPRPEMN 417  
OY 409 SRPLLVRKGTN--FTHLVADRVTGIGDATTYVLFITGDMILKAVSL--GPNVH---L 460  
Db 418 NRPVITKDVYKQFTQIYVDVDAEDG-QYDMFICITGVYVLAVYSIPKFTWTDLEVL 476  
OY 461 IEEQLFQDEP--MRSYLVSSQSKLLFAGSRSQLVQLPVADCIKY-RSCADCVIARDPYC 517  
Db 477 LEEVTVF-REPTAISAMELSTKQOOLYIGSTAGVAQLPHRCDLYGKACACCLARDPYC 535  
OY 518 AM 519  
Db 536 AM 537

## RESULT 5

US-08-835-268-54  
; Sequence 54, Application US/08835268  
; Patent No. 5807826  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey S.  
; APPLICANT: Kolodkin, Alex L.  
; APPLICANT: Matthes, David  
; APPLICANT: Bentley, David R.  
; APPLICANT: O'Connor, Timothy  
; TITLE OF INVENTION: The Semaphorin Gene Family  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 Bush Street, Suite 3200  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/835,268  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,713  
; FILING DATE: 13-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: B94-002-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415)343-4341  
; TELEFAX: (415)343-4342  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 771 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-835-268-54

Query Match 29.0%; Score 932.5; DB 1: Length 771;  
Best Local Similarity 38.4%; Pred. No. 4.6e-94;  
Matches 208; Conservative 94; Mismatches 203; Indels 37; Gaps 15;

OY 8 WL--LAARLWGLGIGAEYWM---NLVPRKTVSSGELATVVRPSQTGI---QDFLTTL 58









```

QY 451 AV-----SLGPNVHLIEIOLFDOEPMR-----SLVLSQSKLLFAGSRNQ 491
:      :      :      :      :      :      :      :      :      :
DB 453 VLPGRSGSGPEPILLEEIDAY--SPARSGKRTAQTARIIGLEIDTEGHLELFAVFSGC 510
:      :      :      :      :      :      :      :      :      :
QY 492 LVQLPVADICKRSC-ADCVLARDPYCAVSNTSR-CVAVGHHFSGSLLO-HVMNSDTSG 548
:      :      :      :      :      :      :      :      :      :
DB 511 IYVLPISRCARHACQRCSCLASDPCYCGW--HSSRCQVDIRGSGTVDVDAQNOESMERG 568
:      :      :      :      :      :      :      :      :      :
QY 549 ICLNLSKISQSGP 561
:      :      :      :      :      :      :      :      :      :
DB 569 DCGDGTGSGQSGP 581
:      :      :      :      :      :      :      :      :      :

RESULT 13
US-08-556-422A-6
: Sequence 6, Application US/08556422A
: Patent No. 6576754
: GENERAL INFORMATION:
: APPLICANT: HALL, Kathryn T.
: APPLICANT: FREEMAN, Gordon J.
: APPLICANT: SCHULTZE, Joachim L.
: APPLICANT: BOUSSIOTIS, Vassiliki
: APPLICANT: NADLER, Lee M.
: TITLE OF INVENTION: NOCULEIC ACIDS ENCODING CD100 MOLECULES
: FILE REFERENCE: DFN-005CPA2
: CURRENT APPLICATION NUMBER: US/08/556,422A
: CURRENT FILING DATE: 1995-11-09
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FASTSEQ for Windows version 4.0
: SEQ ID NO: 6
: LENGTH: 295
: TYPE: PRT
: ORGANISM: Gallus gallus
US-08-556-422A-6

Query Match      22.4%; Score 719.5; DB 4; Length 295;
Best Local Similarity 48.5%; Pred. No. 4,2e-71;
Matches 143; Conservative 49; Mismatches 96; Indels 7; Gaps 4;

QY 232 DDKYFFERRAVESDCYAEQYVAVARVCCKGDMGARTLQRRKTTFLKARLACSAFPMQ 291
:      :      :      :      :      :      :      :      :      :
DB 1 DDKIFFEFVEVVEREFCKLMIPRIARCKRDGGLRLQKMTSFLKARLICITIPDKN 60
:      :      :      :      :      :      :      :      :      :
QY 232 LYFNLOAMHTLQDTSMINTTFEYVFOAGMDMTLSAICEYOLEEIQRYE-EGPYKE--- 347
:      :      :      :      :      :      :      :      :      :
DB 61 LFIENINVFETLSKSTLKEPVYGVFTPOLNNVGLSACAVNMLSAVEEFSGKYMGSAT 120
:      :      :      :      :      :      :      :      :      :
QY 348 YHEAQKDRYTDPPSPRPGSCINNMHRRHGYSLSLELDNLLNPFVKKHPIMEQVGR 407
:      :      :      :      :      :      :      :      :      :
DB 121 VEQSHTKMYRNGELIPNRPFGACINNBARALNTVSLNLPDKTLOFVXDHPIMDSVTPV 180
:      :      :      :      :      :      :      :      :      :
QY 408 WSRPLLVKGTFTFLVADRVTLGDATYTVLFTGTGDMILKAVSLGPNVHLIEIOLF 467
:      :      :      :      :      :      :      :      :      :
DB 161 GDRPLRLVADRVKTYTQIVDVRAALNGITYDVWFISTDGCALHKAISTENGMIHIEETOLF 240
:      :      :      :      :      :      :      :      :      :
QY 468 DQ-EPKRSILVLS--QSKLLFAGSRSLQVLPVADCIKRRSCADCVLARDPYCAM 519
:      :      :      :      :      :      :      :      :      :
DB 241 PFEEVYQTLILSSKSRRLVLAGSNSGVSPVACFDYTTTCFDCVLARDPYCAM 295
:      :      :      :      :      :      :      :      :      :

RESULT 14
US-08-121-713D-58
: Sequence 58, Application US/08121713D
: Patent No. 5639856
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.
: APPLICANT: Matthes, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: The Semaphorin Gene Family
: NUMBER OF SEQUENCES: 100

```

```

: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 Bush Street, Suite 3200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/121,713D
: FILING DATE: 13-SEP-1993
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: B94-002-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)343-4341
: TELEFAX: (415) 343-4342
: TELEX:
: INFORMATION FOR SEQ ID NO: 58:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 730 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-121-713D-58

Query Match      21.9%; Score 704.5; DB 1; Length 730;
Best Local Similarity 31.3%; Pred. No. 9,2e-69;
Matches 190; Conservative 107; Mismatches 245; Indels 65; Gaps 21;

```

```

QY 6 AVMILAARLWGIGICGEVMMNLV-PRKTVSGELATVRRFSOTGIQDTLTLTEPTGL 64
:      :      :      :      :      :      :      :      :      :
DB 4 ALVAVALLM-VALLAAVAVNDVSPKRYVQFGE-ERVQHLNESHKDFKLEKHNSTL 61
:      :      :      :      :      :      :      :      :      :
QY 65 LVYGAREALFASMEALE--LOGAISMEDAPVEKTECIQKGNKNOTECFNFTRFLQPYNA 122
:      :      :      :      :      :      :      :      :      :
DB 62 L-VGARANIYVNNISLRDLTEFTEDRIEMHSSGAHRELCYKGR-SEDDCQNTIRVLAKIDD 119
:      :      :      :      :      :      :      :      :      :
QY 123 SHLYVGYAFAOPKCTYVNMLTFTLEHGEF-----EDGKGPYDPAKAGILVDGELY 177
:      :      :      :      :      :      :      :      :      :
DB 120 DRYLCGTNAKYPRLCRH-----YALKDGDYVEKEVREGGLCPDFDHNSTAIYSGQLY 174
:      :      :      :      :      :      :      :      :      :
QY 178 SATLNFLGTEPTILRNKGRPHISMKTETLAFMLNEPHFVGSAYVPSVGSFTGDDKXVF 237
:      :      :      :      :      :      :      :      :      :
DB 175 SATVADFSGTDLIYR--GPLRTERSDLNQ--LMAPNFVNTMEY-----NDFTIF 220
:      :      :      :      :      :      :      :      :      :
QY 238 FFRERAVESDCYAEQYVAVARVCCKGDMGARTLQRRKTTFLKARLACSAF-NMOLYENO 296
:      :      :      :      :      :      :      :      :      :
DB 221 FFRERAVETINCKAITYSHAVARVCKHDKGPHOGGDMKTSFLKSLRNCVPGDYPTFVNE 280
:      :      :      :      :      :      :      :      :      :
QY 297 LQAMHTLQDTSMINTT---FFGVFOAGMDMTLSAICEYOLEEIQRYEFGPYKEHIEAQ 353
:      :      :      :      :      :      :      :      :      :
DB 281 IGSTDIIEGNAGGVYEKITYGVFTTPVNSIGSNAVCAFSKMSILESFGPKRKEQETMS 340
:      :      :      :      :      :      :      :      :      :
QY 354 KMDRYTD-PVPSRPGSCINNMHRRHGYSLSLELDNLLNPFVKKHPIMEQVGRPKMSRPL 412
:      :      :      :      :      :      :      :      :      :
DB 341 NMLAVPSLKVPEPRPGQVND-----SRLPDVSNVFPKSHTLMDQVAPAFETRTI 391
:      :      :      :      :      :      :      :      :      :
QY 413 LVKKGST--NFTHLVAD-RVYGLDGANTYTVLFTGTGDMILKAVSLGPNVH-----LIE 462
:      :      :      :      :      :      :      :      :      :
DB 392 LIRISLQYRETKIADVQVQRTPDGKAYVDLFTGTGDMIKALNSASFSSDPTVDSVYLE 451
:      :      :      :      :      :      :      :      :      :
QY 463 ELQTFDQ-EPKRSILVLS-----QSKLLFAGSRSLQVLPVADCIKRRSCADCVLARDPY 516
:      :      :      :      :      :      :      :      :      :
DB 453 ELQVLPKGVPAKLVVVRNDGDSKLVVYSDEILAIAIKRHGSGDKITKRCRCVSLQDPY 511
:      :      :      :      :      :      :      :      :      :
QY 517 CAMSVNTSRCAVAVGHFQSL---LIQHVMTSDTSGICMLNSKSIQSGPXPKNITVYVAGT 572
:      :      :      :      :      :      :      :      :      :

```



Db 512 CAMDNEVEKCTAVGSPDMSAGKRPTIONISLGEHKACGGRPOTETIVASPVPTOPTTKSSG 571  
QY 573 DLVLPCH 579  
Db 572 DPHVSIH 578

RESULT 15  
US-08-835-268-58

Sequence 58, Application US/08835268  
Patent No. 5607826  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Kolodkin, Alex L.  
APPLICANT: Matthes, David R.  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,268  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 730 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-835-268-58

Query Match 21.9%, Score 704.5; DB 1; Length 730;  
Best Local Similarity 31.3%; Pred. No. 9.2e-69;  
Matches 190; Conservative 107; Mismatches 245; Indels 65; Gaps 21;

QY 6 AVMLAARLMGLGAGFVWMLV-PRKTVSSGELATVVRRSQIGIDFLITLLEPTGL 64  
Db 4 ALVAVALLW-VALHAAWVNDVSPKNTVQGE-ERVQRLGNESHKDFKLEKDNHSL 61  
QY 65 LVGAREALFAFSMEALE--LQGAISMEAPVEKTECIQKGNKNOTECFNFIRFLQPYNA 122  
Db 62 L-VGARNTVYINISRLDLEFTEQRTEHMSGAHRELCTLKGR-SEDDCONTIRYIAKIID 119  
QY 123 SHLYVCGTYAPQPKCTYVNMULTLEHGEF-----EDGKGRCPYDPAKGAGLLVDEGLY 177  
Db 120 DRVLICGTNAYKPLCRH-----YALKDGDYVEKEYERGLCPDPDHNSTAIYSEGOLY 174  
QY 178 SATLNNPGLTEPIILARNNGPHHSKTEYLAFWLNPHVGSATYVESVGSFTGDDKXYF 237

Db 175 SATVADSESGTDPPLIYR--GPIATERSDIKQ--LNAPNFVNTMEY-----NDLTFE 220  
QY 238 FFERAVASDCYAEQVAVARAVAVKCGMDGARTLQKWTFTLKARLACASAP-NNQLYENO 296  
Db 221 FFERAVAYINCGRATYRVARAVCKHDKGGRHOGDRNTSLKSLNCSVYGDIPFYENE 280  
QY 297 LQAMETLQDTSMHNT---FFGVFOAGMDYLSAICEYOLEEIQREVEGPKYKEHEBAQ 353  
Db 281 IQSTSDILEGAVNGOVEKLIYGVFTTPVNSIGSAVCAFSKMSLIESPDGPFKBOETMNS 340  
QY 354 KWDRTD-PVSPRPGSCINNMHRRHGTSSLELPDLNLNVKKHPIEMEDYGRPMRSL 412  
Db 341 NMLAVPSLKVPEPRGQCVD-----SRTLPDVSVNFVKSHTLMDAEPAPAFETRP 391  
QY 413 LVKKGTT--NFTHLVAD-RVTDGAGTYVLFITGDDGLLRAVSLGPMVH-----LIE 462  
Db 392 LIRISLOYRFTKRIANDQAVRTPDGKAYVLFITGDDGKAYITALNSASDSDPYDVYIE 451  
QY 463 ELQLFDQ-EPMRSLVL-----SQSKLLFAGSRSLQVLPVADCIKYSACDCLYARDPY 516  
Db 452 ELQVLPQGVVPVKNLTVVRMDGDDSKLVVYSDDEILAIKLRHCGSDKITYNCRECVSLQDPY 511  
QY 517 CAMSVNTRCVAVAGHFGSL----LIDHVMTSPTSGICNLGSKIQSGPXPKNITVAVGT 572  
Db 512 CAMDNEVEKCTAVGSPDMSAGKRPTIONISLGEHKACGGRPOTETIVASPVPTOPTTKSSG 571  
QY 573 DLVLPCH 579  
Db 572 DPHVSIH 578

Search completed: August 7, 2003, 13:51:21  
Job time : 31 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 2003, 13:50:00 ; Search time 54 Seconds

(without alignments)  
1310.756 Million cell updates/sec

Title: US-10-002-050-14

Perfect score: 3219  
Sequence: 1 MAPHAWMLAARLWGIG.....PCHLSLALPDSPRESSV 596

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/FCI\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/FCIS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3217	99.9	596	14	US-10-003-152-14 Sequence 14, Appl
2	3217	99.9	596	15	US-10-002-050-14 Sequence 14, Appl
3	3217	99.9	596	15	US-10-002-304-14 Sequence 14, Appl
4	3115.5	96.8	590	14	US-10-003-152-12 Sequence 12, Appl
5	3115.5	96.8	590	15	US-10-002-050-12 Sequence 12, Appl
6	3115.5	96.8	590	15	US-10-002-304-12 Sequence 12, Appl
7	3113.5	96.7	833	15	US-10-149-819-4 Sequence 22, Appl
8	3107.5	96.5	624	14	US-10-003-152-22 Sequence 22, Appl
9	3107.5	96.5	624	15	US-10-002-050-22 Sequence 22, Appl
10	3107.5	96.5	624	15	US-10-002-304-22 Sequence 22, Appl
11	2720	84.5	834	15	US-09-759-1308-335 Sequence 35, Appl
12	2720	84.5	834	15	US-10-189-123-65 Sequence 218, App
13	2635	81.9	785	10	US-09-989-920-218 Sequence 253, App
14	1240	38.5	837	11	US-09-946-374-253 Sequence 454, App
15	1240	38.5	837	14	US-10-052-586-454 Sequence 454, App

16	1240	38.5	837	15	US-10-174-590-454	Sequence 454, App
17	1240	38.5	837	15	US-10-176-758-454	Sequence 454, App
18	1240	38.5	837	15	US-10-175-737-454	Sequence 454, App
19	1240	38.5	837	15	US-10-173-706-454	Sequence 454, App
20	1240	38.5	837	15	US-10-175-738-454	Sequence 454, App
21	1240	38.5	837	15	US-10-175-752-454	Sequence 454, App
22	1240	38.5	837	15	US-10-176-482-454	Sequence 454, App
23	1240	38.5	837	15	US-10-176-757-454	Sequence 454, App
24	1240	38.5	837	15	US-10-176-913-454	Sequence 454, App
25	1240	38.5	837	15	US-10-180-552-454	Sequence 454, App
26	1240	38.5	837	15	US-10-180-557-454	Sequence 454, App
27	1240	38.5	837	15	US-10-173-700-454	Sequence 454, App
28	1240	38.5	837	15	US-10-174-572-454	Sequence 454, App
29	1240	38.5	837	15	US-10-174-579-454	Sequence 454, App
30	1240	38.5	837	15	US-10-174-582-454	Sequence 454, App
31	1240	38.5	837	15	US-10-174-588-454	Sequence 454, App
32	1240	38.5	837	15	US-10-175-739-454	Sequence 454, App
33	1240	38.5	837	15	US-10-175-740-454	Sequence 454, App
34	1240	38.5	837	15	US-10-175-743-454	Sequence 454, App
35	1240	38.5	837	15	US-10-176-488-454	Sequence 454, App
36	1240	38.5	837	15	US-10-176-492-454	Sequence 454, App
37	1240	38.5	837	15	US-10-176-747-454	Sequence 454, App
38	1240	38.5	837	15	US-10-176-750-454	Sequence 454, App
39	1240	38.5	837	15	US-10-176-985-454	Sequence 454, App
40	1240	38.5	837	15	US-10-176-987-454	Sequence 454, App
41	1240	38.5	837	15	US-10-176-992-454	Sequence 454, App
42	1240	38.5	837	15	US-10-176-993-454	Sequence 454, App
43	1240	38.5	837	15	US-10-184-658-454	Sequence 454, App
44	1240	38.5	837	15	US-10-176-991-454	Sequence 454, App
45	1240	38.5	837	15	US-10-173-695-454	Sequence 454, App

#### ALIGNMENTS

RESULT 1  
US-10-003-152-14  
Sequence 14, Application US/10003152  
Publication No. US20020151494A1  
GENERAL INFORMATION:  
APPLICANT: Shinketsu, Richard  
APPLICANT: Fernandes, Elma  
APPLICANT: Vermet, Corine  
APPLICANT: Yang, Meljia  
APPLICANT: Boldog, Ferenc  
APPLICANT: Herrmann, John  
TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin  
FILE REFERENCE: 15966-554 Cura-54 CON-512  
CURRENT FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: 09/604,286  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/140,584  
PRIOR FILING DATE: 1999-06-23  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 596  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: variation  
LOCATION: (562)  
OTHER INFORMATION: Xaa may be any amino acid  
US-10-003-152-14

Query Match 99.9%; Score 3217; DB 14; Length 596;  
Best Local Similarity 100.0%; Pred. No. 4e-318;

Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPHAWMLAARLWGIGAEVMMNVPRKTVSGSELATVVRRSQIGIDFLTLTLE 60  
DB 1 MAPHAWMLAARLWGIGAEVMMNVPRKTVSGSELATVVRRSQIGIDFLTLTLE 60

```
OY 61 PGLLVYVAGREALFAFSMSELELOGAIISWEAPVEKTECIOGKNNQTECFNFIIRLOPY 120
DB 61 PGLLVYVAGREALFAFSMSELELOGAIISWEAPVEKTECIOGKNNQTECFNFIIRLOPY 120
OY 121 NASHLYVCGTYAFOPRCCTVNNMLFTFLEHGEFEDGKCPYDPAKHAGLVDGELYSAT 180
DB 121 NASHLYVCGTYAFOPRCCTVNNMLFTFLEHGEFEDGKCPYDPAKHAGLVDGELYSAT 180
OY 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPEVSGSFTGDDDKYFFFR 240
DB 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPEVSGSFTGDDDKYFFFR 240
OY 241 ERAVESDCYAEQVAVAVARCKGDMGARTLOKRWTFPLKARLACAPMMOLYFNLOAM 300
DB 241 ERAVESDCYAEQVAVAVARCKGDMGARTLOKRWTFPLKARLACAPMMOLYFNLOAM 300
OY 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQPYREYHEEAKMDRYTD 360
DB 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQPYREYHEEAKMDRYTD 360
OY 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFVKKHPLMEBOVGPRMSRPLLKKGTFN 420
DB 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFVKKHPLMEBOVGPRMSRPLLKKGTFN 420
OY 421 THLVADRVGTGLDCAITTVLPIGTGDMVLKAVSLGFWHLIEBQLFDDPMSRYLVSOS 480
DB 421 THLVADRVGTGLDCAITTVLPIGTGDMVLKAVSLGFWHLIEBQLFDDPMSRYLVSOS 480
OY 481 KLLFAGSRSQLVOLPVADCIKRYSCADCVLADPFCAMSVNTSRCVAVGGHGSLLIOH 540
DB 481 KLLFAGSRSQLVOLPVADCIKRYSCADCVLADPFCAMSVNTSRCVAVGGHGSLLIOH 540
OY 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596
DB 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596

RESULT 2
US-10-002-050-14
/ Sequence 14, Application US/10002050
/ Publication No. US20030032095A1
/ GENERAL INFORMATION:
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Vernet, Corline
/ APPLICANT: Yang, Meljia
/ APPLICANT: Boldog, Ferenc
/ APPLICANT: Hermann, John
/ TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Semaph
/ FILE REFERENCE: 15966-554 Cura-54 CON-S14
/ CURRENT APPLICATION NUMBER: US/10/002,050
/ PRIOR FILING DATE: 2001-11-02
/ PRIOR APPLICATION NUMBER: 09/604,286
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 60/140,584
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 596
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: variation
/ LOCATION: (562)
/ OTHER INFORMATION: Xaa may be any amino acid
US-10-002-050-14

Query Match 99.9%; Score 3217; DB 15; Length 596;
Best Local Similarity 100.0%; Pred. No. 4e-318;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MAPHNAVMLLAARLWGLGIGAEVWNLVPRKTVSSSELATVVRFSQGTIGDFTLTUTE 60
DB 1 MAPHNAVMLLAARLWGLGIGAEVWNLVPRKTVSSSELATVVRFSQGTIGDFTLTUTE 60
OY 61 PGLLVYVAGREALFAFSMSELELOGAIISWEAPVEKTECIOGKNNQTECFNFIIRLOPY 120
DB 61 PGLLVYVAGREALFAFSMSELELOGAIISWEAPVEKTECIOGKNNQTECFNFIIRLOPY 120
OY 121 NASHLYVCGTYAFOPRCCTVNNMLFTFLEHGEFEDGKCPYDPAKHAGLVDGELYSAT 180
DB 121 NASHLYVCGTYAFOPRCCTVNNMLFTFLEHGEFEDGKCPYDPAKHAGLVDGELYSAT 180
OY 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPEVSGSFTGDDDKYFFFR 240
DB 181 LNNEFLCTEPIILRNMGPHHSMTETYLAFMLNEPHFVGSAYVPEVSGSFTGDDDKYFFFR 240
OY 241 ERAVESDCYAEQVAVAVARCKGDMGARTLOKRWTFPLKARLACAPMMOLYFNLOAM 300
DB 241 ERAVESDCYAEQVAVAVARCKGDMGARTLOKRWTFPLKARLACAPMMOLYFNLOAM 300
OY 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQPYREYHEEAKMDRYTD 360
DB 301 HTLQDTSMTNTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQPYREYHEEAKMDRYTD 360
OY 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFVKKHPLMEBOVGPRMSRPLLKKGTFN 420
DB 361 PVPSPRPGSCINNMHRRHGYTSSLELPDNLNFVKKHPLMEBOVGPRMSRPLLKKGTFN 420
OY 421 THLVADRVGTGLDCAITTVLPIGTGDMVLKAVSLGFWHLIEBQLFDDPMSRYLVSOS 480
DB 421 THLVADRVGTGLDCAITTVLPIGTGDMVLKAVSLGFWHLIEBQLFDDPMSRYLVSOS 480
OY 481 KLLFAGSRSQLVOLPVADCIKRYSCADCVLADPFCAMSVNTSRCVAVGGHGSLLIOH 540
DB 481 KLLFAGSRSQLVOLPVADCIKRYSCADCVLADPFCAMSVNTSRCVAVGGHGSLLIOH 540
OY 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596
DB 541 VMTSDTSGICNLGSKIQSGPXPKNITVAGTDLVLPCHLSSNLALPDSNPESSV 596

RESULT 3
US-10-002-304-14
/ Sequence 14, Application US/10002304
/ Publication No. US20030036185A1
/ GENERAL INFORMATION:
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Vernet, Corline
/ APPLICANT: Yang, Meljia
/ APPLICANT: Boldog, Ferenc
/ APPLICANT: Hermann, John
/ TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
/ FILE REFERENCE: 15966-554 Cura-54 CON-S8
/ CURRENT APPLICATION NUMBER: US/10/002,304
/ PRIOR FILING DATE: 2001-11-02
/ PRIOR APPLICATION NUMBER: 09/604,286
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 60/140,584
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 596
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: variation
/ LOCATION: (562)
/ OTHER INFORMATION: Xaa may be any amino acid
US-10-002-304-14

Query Match 99.9%; Score 3217; DB 15; Length 596;
```

Best Local Similarity 100.0%; Pred. No. 4e-318;  
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAPHAAVWLLAARLMGIGIGAEVMMNLVPRKTVSSGELATVVRPFSOTGIDFLLTLTE 60
    1 MAPHAAVWLLAARLMGIGIGAEVMMNLVPRKTVSSGELATVVRPFSOTGIDFLLTLTE 60
Db 1 MAPHAAVWLLAARLMGIGIGAEVMMNLVPRKTVSSGELATVVRPFSOTGIDFLLTLTE 60
QY 61 PTGLLYGARBALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRIFLOPY 120
    61 PTGLLYGARBALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRIFLOPY 120
Db 61 PTGLLYGARBALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRIFLOPY 120
QY 121 NASHLYVCGTAAPOPKCTVYVNMFLFTLEHGEFEGKGCPRDPKAGAGLLVDELVSAT 180
    121 NASHLYVCGTAAPOPKCTVYVNMFLFTLEHGEFEGKGCPRDPKAGAGLLVDELVSAT 180
Db 121 NASHLYVCGTAAPOPKCTVYVNMFLFTLEHGEFEGKGCPRDPKAGAGLLVDELVSAT 180
QY 181 LNNFLGTEPILIRNMGPHHSKTEYLAFWLNPEHFVSAAVYPSVGSFTGDDDVYFFER 240
    181 LNNFLGTEPILIRNMGPHHSKTEYLAFWLNPEHFVSAAVYPSVGSFTGDDDVYFFER 240
Db 181 LNNFLGTEPILIRNMGPHHSKTEYLAFWLNPEHFVSAAVYPSVGSFTGDDDVYFFER 240
QY 241 ERAVESDCYAEQVAVARAVAVKCGDMGARTLQRRMTFLKARLACSAAPNMQLYFNQLOAM 300
    241 ERAVESDCYAEQVAVARAVAVKCGDMGARTLQRRMTFLKARLACSAAPNMQLYFNQLOAM 300
Db 241 ERAVESDCYAEQVAVARAVAVKCGDMGARTLQRRMTFLKARLACSAAPNMQLYFNQLOAM 300
QY 301 HTLQDTSNHNTEFFGVQAOQMDMYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRYTD 360
    301 HTLQDTSNHNTEFFGVQAOQMDMYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRYTD 360
Db 301 HTLQDTSNHNTEFFGVQAOQMDMYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRYTD 360
QY 361 PVSPRPGSCINNNHRRHGYTSSLELDPNLIINFVKHPLMEQVGRPMRSLYLKKGTFN 420
    361 PVSPRPGSCINNNHRRHGYTSSLELDPNLIINFVKHPLMEQVGRPMRSLYLKKGTFN 420
Db 361 PVSPRPGSCINNNHRRHGYTSSLELDPNLIINFVKHPLMEQVGRPMRSLYLKKGTFN 420
QY 421 THLVADRYTGLDGAATYVLFITGQDGLKAVSLGAPWVHLIEELQLEFQDEPMRSLVLSQS 480
    421 THLVADRYTGLDGAATYVLFITGQDGLKAVSLGAPWVHLIEELQLEFQDEPMRSLVLSQS 480
Db 421 THLVADRYTGLDGAATYVLFITGQDGLKAVSLGAPWVHLIEELQLEFQDEPMRSLVLSQS 480
QY 481 KKLIFAGRSQVQLPVADCIKYSRSCADCVLARDPYCAMSNTSRCAVAGHFSLLIOH 540
    481 KKLIFAGRSQVQLPVADCIKYSRSCADCVLARDPYCAMSNTSRCAVAGHFSLLIOH 540
Db 481 KKLIFAGRSQVQLPVADCIKYSRSCADCVLARDPYCAMSNTSRCAVAGHFSLLIOH 540
QY 541 VMTSDTSGICMLRSGSKIOSGPKPNITVVAAGTDVLPCHLSSNLALPDSNPESSV 596
    541 VMTSDTSGICMLRSGSKIOSGPKPNITVVAAGTDVLPCHLSSNLALPDSNPESSV 596
Db 541 VMTSDTSGICMLRSGSKIOSGPKPNITVVAAGTDVLPCHLSSNLALPDSNPESSV 596
```

## RESULT 4

US-10-003-152-12  
; Sequence 12, Application US/10003152  
; Publication No. US20020151494A1  
; GENERAL INFORMATION:  
; APPLICANT: Shinkels, Richard  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Vernet, Corine  
; APPLICANT: Yang, Meijia  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Herrmann, John  
; TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin-7  
; FILE REFERENCE: 15966-554 Cura-54 CON-S12  
; CURRENT APPLICATION NUMBER: US/10/003.152  
; PRIOR FILING DATE: 2001-11-02  
; PRIOR APPLICATION NUMBER: 09/604,286  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,584  
; PRIOR FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-003-152-12

Query Match 96.8%; Score 3115.5; DB 14; Length 590;  
Best Local Similarity 99.1%; Pred. No. 8.6e-308;

Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

```
QY 1 MAPHAAVWLLAARLMGIGIGAEVMMNLVPRKTVSSGELATVVRPFSOTGIDFLLTLTE 60
    1 MAPHAAVWLLAARLMGIGIGAEVMMNLVPRKTVSSGELATVVRPFSOTGIDFLLTLTE 60
Db 1 MAPHAAVWLLAARLMGIGIGAEVMMNLVPRKTVSSGELATVVRPFSOTGIDFLLTLTE 60
QY 61 PTGLLYGARBALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRIFLOPY 120
    61 PTGLLYGARBALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRIFLOPY 120
Db 61 PTGLLYGARBALFAFSMEALELQGAISWEAPVEKTECIQKGNNOECFNFRIFLOPY 120
QY 121 NASHLYVCGTAAPOPKCTVYVNMFLFTLEHGEFEGKGCPRDPKAGAGLLVDELVSAT 180
    121 NASHLYVCGTAAPOPKCTVYVNMFLFTLEHGEFEGKGCPRDPKAGAGLLVDELVSAT 180
Db 121 NASHLYVCGTAAPOPKCTVYVNMFLFTLEHGEFEGKGCPRDPKAGAGLLVDELVSAT 180
QY 181 LNNFLGTEPILIRNMGPHHSKTEYLAFWLNPEHFVSAAVYPSVGSFTGDDDVYFFER 240
    181 LNNFLGTEPILIRNMGPHHSKTEYLAFWLNPEHFVSAAVYPSVGSFTGDDDVYFFER 240
Db 181 LNNFLGTEPILIRNMGPHHSKTEYLAFWLNPEHFVSAAVYPSVGSFTGDDDVYFFER 240
QY 241 ERAVESDCYAEQVAVARAVAVKCGDMGARTLQRRMTFLKARLACSAAPNMQLYFNQLOAM 300
    241 ERAVESDCYAEQVAVARAVAVKCGDMGARTLQRRMTFLKARLACSAAPNMQLYFNQLOAM 300
Db 241 ERAVESDCYAEQVAVARAVAVKCGDMGARTLQRRMTFLKARLACSAAPNMQLYFNQLOAM 300
QY 301 HTLQDTSNHNTEFFGVQAOQMDMYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRYTD 360
    301 HTLQDTSNHNTEFFGVQAOQMDMYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRYTD 360
Db 301 HTLQDTSNHNTEFFGVQAOQMDMYLSAICGYOLEEIOREVEGPKYKEYHEBAQKMDRYTD 360
QY 361 PVSPRPGSCINNNHRRHGYTSSLELDPNLIINFVKHPLMEQVGRPMRSLYLKKGTFN 420
    361 PVSPRPGSCINNNHRRHGYTSSLELDPNLIINFVKHPLMEQVGRPMRSLYLKKGTFN 420
Db 361 PVSPRPGSCINNNHRRHGYTSSLELDPNLIINFVKHPLMEQVGRPMRSLYLKKGTFN 420
QY 421 THLVADRYTGLDGAATYVLFITGQDGLKAVSLGAPWVHLIEELQLEFQDEPMRSLVLSQS 480
    421 THLVADRYTGLDGAATYVLFITGQDGLKAVSLGAPWVHLIEELQLEFQDEPMRSLVLSQS 480
Db 421 THLVADRYTGLDGAATYVLFITGQDGLKAVSLGAPWVHLIEELQLEFQDEPMRSLVLSQS 480
QY 481 KKLIFAGRSQVQLPVADCIKYSRSCADCVLARDPYCAMSNTSRCAVAGHFSLLIOH 540
    481 KKLIFAGRSQVQLPVADCIKYSRSCADCVLARDPYCAMSNTSRCAVAGHFSLLIOH 540
Db 481 KKLIFAGRSQVQLPVADCIKYSRSCADCVLARDPYCAMSNTSRCAVAGHFSLLIOH 540
QY 541 VMTSDTSGICMLRSGSKIOSGPKPNITVVAAGTDVLPCHLSS 582
    541 VMTSDTSGICMLRSGSKIOSGPKPNITVVAAGTDVLPCHLSS 582
Db 541 VMTSDTSGICMLRSGSKIOSGPKPNITVVAAGTDVLPCHLSS 581
```

## RESULT 5

US-10-002-050-12  
; Sequence 12, Application US/10002050  
; Publication No. US20030032095A1  
; GENERAL INFORMATION:  
; APPLICANT: Shinkels, Richard  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Vernet, Corine  
; APPLICANT: Yang, Meijia  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Herrmann, John  
; TITLE OF INVENTION: No. US20030032095A1 Nucleic Acid Sequences Encoding Human Se  
; FILE REFERENCE: 15966-554 Cura-54 CON-S14  
; CURRENT APPLICATION NUMBER: US/10/002.050  
; PRIOR FILING DATE: 2001-11-02  
; PRIOR APPLICATION NUMBER: 09/604,286  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,584  
; PRIOR FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-002-050-12

Query Match 96.8%; Score 3115.5; DB 15; Length 590;  
Best Local Similarity 99.1%; Pred. No. 8.6e-308;  
Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

```

QY 1 MAPHAAVWLLAARLMLGIGIAEYWNLVPRKTVSSSEGLATVVRRESQTSIQDFLTLLTE 60
    1 MAPHAAVWLLAARLMLGIGIAEYWNLVPRKTVSSSEGLATVVRRESQTSIQDFLTLLTE 60
Db 1 MAPHAAVWLLAARLMLGIGIAEYWNLVPRKTVSSSEGLATVVRRESQTSIQDFLTLLTE 60
QY 61 PTGLLYVGARREALFAFMSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQY 120
    61 PTGLLYVGARREALFAFMSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQY 120
Db 61 PTGLLYVGARREALFAFMSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQY 120
QY 121 NASHLYVCGTYAFOPACTYVNMVLTFLHGEFEEDGKCPYPDPAKHAGLLVDGELYSAT 180
    121 NASHLYVCGTYAFOPACTYVNMVLTFLHGEFEEDGKCPYPDPAKHAGLLVDGELYSAT 180
Db 121 NASHLYVCGTYAFOPACTYVNMVLTFLHGEFEEDGKCPYPDPAKHAGLLVDGELYSAT 180
QY 181 LNMFLOTBPIILRNKMPHHSMTETYLAFVLENEPHFVGSAYVPSVGSFTGDDDKYFFFR 240
    181 LNMFLOTBPIILRNKMPHHSMTETYLAFVLENEPHFVGSAYVPSVGSFTGDDDKYFFFR 240
Db 181 LNMFLOTBPIILRNKMPHHSMTETYLAFVLENEPHFVGSAYVPSVGSFTGDDDKYFFFR 240
QY 241 ERAVESDCAEVOVAVARVCKGDMGARTLQKMTTFLKARLACAPMMOLYFNOLQM 300
    241 ERAVESDCAEVOVAVARVCKGDMGARTLQKMTTFLKARLACAPMMOLYFNOLQM 300
Db 241 ERAVESDCAEVOVAVARVCKGDMGARTLQKMTTFLKARLACAPMMOLYFNOLQM 300
QY 301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYQLEIEIQVFEQPYKEYEHEAQKMDRYD 360
    301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYQLEIEIQVFEQPYKEYEHEAQKMDRYD 360
Db 301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYQLEIEIQVFEQPYKEYEHEAQKMDRYD 360
QY 361 PVPSPRPGSCINNMHRRHGITSLSLELPDNLNFKVKKHPLMEBOYGRPMRPLLYKKGTFN 420
    361 PVPSPRPGSCINNMHRRHGITSLSLELPDNLNFKVKKHPLMEBOYGRPMRPLLYKKGTFN 420
Db 361 PVPSPRPGSCINNMHRRHGITSLSLELPDNLNFKVKKHPLMEBOYGRPMRPLLYKKGTFN 420
QY 421 THLVADRTVGLDCAATTVYFIFGTGDMVLLKAVSLGPMVHLIELOLFDDEPMRSLVLSOS 480
    421 THLVADRTVGLDCAATTVYFIFGTGDMVLLKAVSLGPMVHLIELOLFDDEPMRSLVLSOS 480
Db 421 THLVADRTVGLDCAATTVYFIFGTGDMVLLKAVSLGPMVHLIELOLFDDEPMRSLVLSOS 480
QY 481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLARDPYCAMS VNTSRCAVAGHGSLLIQH 540
    481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLARDPYCAMS VNTSRCAVAGHGSLLIQH 540
Db 481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLARDPYCAMS VNTSRCAVAGHGSLLIQH 540
QY 541 VMTSDTSGICNLGRSKIQSGPYKNTTVVAGTDVLPCHLSS 582
    541 VMTSDTSGICNLGRSKIQSGPYKNTTVVAGTDVLPCHLSS 582
Db 541 VMTSDTSGICNLGRSKIQSGPYKNTTVVAGTDVLPCHLSS 582

RESULT 6
US-10-002-304-12
; Sequence 12, Application US/10002304
; Publication No. US20030036185A1
; GENERAL INFORMATION:
; APPLICANT: Shinketsu, Richard
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Yang, Meijia
; APPLICANT: Boldog, Ferenc
; APPLICANT: Heitmann, John
; FILE REFERENCE: 15966-554 Cura-54 CON-58
; CURRENT APPLICATION NUMBER: US/10/002,304
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 09/604,286
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,584
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-304-12

Query Match          96.8%; Score 3115.5; DB 15; Length 590;
Best Local Similarity 99.1%; Pred. No. 8,6e-308;
Matches 577; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

```

```

QY 1 MAPHAAVWLLAARLMLGIGIAEYWNLVPRKTVSSSEGLATVVRRESQTSIQDFLTLLTE 60
    1 MAPHAAVWLLAARLMLGIGIAEYWNLVPRKTVSSSEGLATVVRRESQTSIQDFLTLLTE 60
Db 1 MAPHAAVWLLAARLMLGIGIAEYWNLVPRKTVSSSEGLATVVRRESQTSIQDFLTLLTE 60
QY 61 PTGLLYVGARREALFAFMSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQY 120
    61 PTGLLYVGARREALFAFMSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQY 120
Db 61 PTGLLYVGARREALFAFMSMELALOGAISWEAPVEKTECIQKKNQTECFNIRFLQY 120
QY 121 NASHLYVCGTYAFOPACTYVNMVLTFLHGEFEEDGKCPYPDPAKHAGLLVDGELYSAT 180
    121 NASHLYVCGTYAFOPACTYVNMVLTFLHGEFEEDGKCPYPDPAKHAGLLVDGELYSAT 180
Db 121 NASHLYVCGTYAFOPACTYVNMVLTFLHGEFEEDGKCPYPDPAKHAGLLVDGELYSAT 180
QY 181 LNMFLOTBPIILRNKMPHHSMTETYLAFVLENEPHFVGSAYVPSVGSFTGDDDKYFFFR 240
    181 LNMFLOTBPIILRNKMPHHSMTETYLAFVLENEPHFVGSAYVPSVGSFTGDDDKYFFFR 240
Db 181 LNMFLOTBPIILRNKMPHHSMTETYLAFVLENEPHFVGSAYVPSVGSFTGDDDKYFFFR 240
QY 241 ERAVESDCAEVOVAVARVCKGDMGARTLQKMTTFLKARLACAPMMOLYFNOLQM 300
    241 ERAVESDCAEVOVAVARVCKGDMGARTLQKMTTFLKARLACAPMMOLYFNOLQM 300
Db 241 ERAVESDCAEVOVAVARVCKGDMGARTLQKMTTFLKARLACAPMMOLYFNOLQM 300
QY 301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYQLEIEIQVFEQPYKEYEHEAQKMDRYD 360
    301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYQLEIEIQVFEQPYKEYEHEAQKMDRYD 360
Db 301 HTLDQTSWNTTFFGVFOAQMGDMYLSAICEYQLEIEIQVFEQPYKEYEHEAQKMDRYD 360
QY 361 PVPSPRPGSCINNMHRRHGITSLSLELPDNLNFKVKKHPLMEBOYGRPMRPLLYKKGTFN 420
    361 PVPSPRPGSCINNMHRRHGITSLSLELPDNLNFKVKKHPLMEBOYGRPMRPLLYKKGTFN 420
Db 361 PVPSPRPGSCINNMHRRHGITSLSLELPDNLNFKVKKHPLMEBOYGRPMRPLLYKKGTFN 420
QY 421 THLVADRTVGLDCAATTVYFIFGTGDMVLLKAVSLGPMVHLIELOLFDDEPMRSLVLSOS 480
    421 THLVADRTVGLDCAATTVYFIFGTGDMVLLKAVSLGPMVHLIELOLFDDEPMRSLVLSOS 480
Db 421 THLVADRTVGLDCAATTVYFIFGTGDMVLLKAVSLGPMVHLIELOLFDDEPMRSLVLSOS 480
QY 481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLARDPYCAMS VNTSRCAVAGHGSLLIQH 540
    481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLARDPYCAMS VNTSRCAVAGHGSLLIQH 540
Db 481 KLLFAGSRSQVLQVLPVADCIKRYSCADCVLARDPYCAMS VNTSRCAVAGHGSLLIQH 540
QY 541 VMTSDTSGICNLGRSKIQSGPYKNTTVVAGTDVLPCHLSS 582
    541 VMTSDTSGICNLGRSKIQSGPYKNTTVVAGTDVLPCHLSS 582
Db 541 VMTSDTSGICNLGRSKIQSGPYKNTTVVAGTDVLPCHLSS 582

RESULT 7
US-10-149-819-4
; Sequence 4, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BUREFORD, Neil
; FILE REFERENCE: PE-0760 PCR
; CURRENT APPLICATION NUMBER: US/10/149,819
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 1674661CD1
US-10-149-819-4

```

Query Match 96.7%; Score 3113.5; DB 15; Length 833;  
 Best Local Similarity 96.3%; Pred. No. 2.3e-307;  
 Matches 579; Conservative 4; Mismatches 7; Indels 11; Gaps 2;

```

QY 1 MAPHAWMLAARLWGLGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTLITE 60
DB 1 MAPHAWMLAARLWGLGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTLITE 60
QY 61 PTGLLYGAREALFAFSMEALELOGAISWEAPVEKTECIQKGNKNOECENFIRFLQPY 120
DB 61 PTGLLYGAREALFAFSMEALELOGAISWEAPVEKTECIQKGNKNOECENFIRFLQPY 120
QY 121 NASHLYVCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRYPAPKAGHGLLVDELYSAT 180
DB 121 NASHLYVCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRYPAPKAGHGLLVDELYSAT 180
QY 181 LNNFLGTEPITLRNMGPHHSKTEYLAFWLNPHFVGSAYVPEVSGFTGDDKYFFFR 240
DB 181 LNNFLGTEPITLRNMGPHHSKTEYLAFWLNPHFVGSAYVPEVSGFTGDDKYFFFR 240
QY 241 ERAVESDCYABOYVARAVARVCKGDMGARTLQKRTTFLKARLACSAPNMQLYFNQLQAM 300
DB 241 ERAVESDCYABOYVARAVARVCKGDMGARTLQKRTTFLKARLACSAPNMQLYFNQLQAM 300
QY 301 HTLQDTSNHNTEFFGVFOQWGDWYLSAICEYOLEIQRVDEGPYKEYHEBAOKMDRTD 360
DB 301 HTLQDTSNHNTEFFGVFOQWGDWYLSAICEYOLEIQRVDEGPYKEYHEBAOKMDRTD 360
QY 361 PVPSPRGSCINNNHRRHGYTSSLELDPNIIINFVKKHPHLMEOVGPRKSRPLVKKGNF 420
DB 361 PVPSPRGSCINNNHRRHGYTSSLELDPNIIINFVKKHPHLMEOVGPRKSRPLVKKGNF 420
QY 421 THLVADRYTGLDGYATYVLTFTGTDGWLKAVSLGPMVHLIEQLDFQEPERSLVLSQS 480
DB 421 THLVADRYTGLDGYATYVLTFTGTDGWLKAVSLGPMVHLIEQLDFQEPERSLVLSQS 480
QY 481 KLLPAGRSQVLPVADCIKRYSCADCYLAADPYCAMSVTSCVAVGSHGSLLIQH 540
DB 481 KLLPAGRSQVLPVADCIKRYSCADCYLAADPYCAMSVTSCVAVGSHGSLLIQH 540
QY 541 VMTSDTSGICNLGSKIOSGPKNTIYVAGTDVLPCHLSSNLA-----LPDSN 590
DB 541 VMTSDTSGICNLGSK-KVRPTPKNTIYVAGTDVLPCHLSSNLAHAWTFGGRDLPAEQ 599
QY 591 P 591
DB 600 P 600

```

RESULT 8  
 US-10-003-152-22  
 : Sequence 22, Application US/10003152  
 : Publication No. US20020151494A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Shimkets, Richard  
 : APPLICANT: Fernandes, Elma  
 : APPLICANT: Vernet, Corine  
 : APPLICANT: Yang, Meljia  
 : APPLICANT: Boldog, Ferenc  
 : APPLICANT: Herrmann, John  
 : TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin-1  
 : FILE REFERENCE: 15966-554 Cura-54 CON-S12  
 : CURRENT APPLICATION NUMBER: US/10/003.152  
 : CURRENT FILING DATE: 2001-11-02  
 : PRIOR APPLICATION NUMBER: 09/604, 286  
 : PRIOR FILING DATE: 2000-06-22  
 : PRIOR APPLICATION NUMBER: 60/140, 584  
 : PRIOR FILING DATE: 1999-06-23  
 : NUMBER OF SEQ ID NOS: 49  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 22  
 : LENGTH: 624

TYPE: PRT  
 : ORGANISM: Homo sapiens  
 US-10-003-152-22

Query Match 96.5%; Score 3107.5; DB 14; Length 624;  
 Best Local Similarity 99.0%; Pred. No. 6.1e-307;  
 Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

```

QY 1 MAPHAWMLAARLWGLGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTLITE 60
DB 1 MAPHAWMLAARLWGLGIGAEVMMNLVPRKTVSSGELATVYRRFSQTGIDDFLTLITE 60
QY 61 PTGLLYGAREALFAFSMEALELOGAISWEAPVEKTECIQKGNKNOECENFIRFLQPY 120
DB 61 PTGLLYGAREALFAFSMEALELOGAISWEAPVEKTECIQKGNKNOECENFIRFLQPY 120
QY 121 NASHLYVCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRYPAPKAGHGLLVDELYSAT 180
DB 121 NASHLYVCGTYAFQPKCTYVNMFLTLEHGEFEDGKGCPRYPAPKAGHGLLVDELYSAT 180
QY 181 LNNFLGTEPITLRNMGPHHSKTEYLAFWLNPHFVGSAYVPEVSGFTGDDKYFFFR 240
DB 181 LNNFLGTEPITLRNMGPHHSKTEYLAFWLNPHFVGSAYVPEVSGFTGDDKYFFFR 240
QY 241 ERAVESDCYABOYVARAVARVCKGDMGARTLQKRTTFLKARLACSAPNMQLYFNQLQAM 300
DB 241 ERAVESDCYABOYVARAVARVCKGDMGARTLQKRTTFLKARLACSAPNMQLYFNQLQAM 300
QY 301 HTLQDTSNHNTEFFGVFOQWGDWYLSAICEYOLEIQRVDEGPYKEYHEBAOKMDRTD 360
DB 301 HTLQDTSNHNTEFFGVFOQWGDWYLSAICEYOLEIQRVDEGPYKEYHEBAOKMDRTD 360
QY 361 PVPSPRGSCINNNHRRHGYTSSLELDPNIIINFVKKHPHLMEOVGPRKSRPLVKKGNF 420
DB 361 PVPSPRGSCINNNHRRHGYTSSLELDPNIIINFVKKHPHLMEOVGPRKSRPLVKKGNF 420
QY 421 THLVADRYTGLDGYATYVLTFTGTDGWLKAVSLGPMVHLIEQLDFQEPERSLVLSQS 480
DB 421 THLVADRYTGLDGYATYVLTFTGTDGWLKAVSLGPMVHLIEQLDFQEPERSLVLSQS 480
QY 481 KLLPAGRSQVLPVADCIKRYSCADCYLAADPYCAMSVTSCVAVGSHGSLLIQH 540
DB 481 KLLPAGRSQVLPVADCIKRYSCADCYLAADPYCAMSVTSCVAVGSHGSLLIQH 540
QY 541 VMTSDTSGICNLGSKIOSGPKNTIYVAGTDVLPCHLSS 582
DB 541 VMTSDTSGICNLGSK-KVRPTPKNTIYVAGTDVLPCHLSS 581

```

RESULT 9  
 US-10-002-050-22  
 : Sequence 22, Application US/10002050  
 : Publication No. US20030032095A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Shimkets, Richard  
 : APPLICANT: Fernandes, Elma  
 : APPLICANT: Vernet, Corine  
 : APPLICANT: Yang, Meljia  
 : APPLICANT: Boldog, Ferenc  
 : APPLICANT: Herrmann, John  
 : TITLE OF INVENTION: No. US20030032095A1 Nucleic Acid Sequences Encoding Human Se  
 : FILE REFERENCE: 15966-554 Cura-54 CON-S14  
 : CURRENT APPLICATION NUMBER: US/10/002.050  
 : CURRENT FILING DATE: 2001-11-02  
 : PRIOR APPLICATION NUMBER: 09/604, 286  
 : PRIOR FILING DATE: 2000-06-22  
 : PRIOR APPLICATION NUMBER: 60/140, 584  
 : PRIOR FILING DATE: 1999-06-23  
 : NUMBER OF SEQ ID NOS: 49  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 22  
 : LENGTH: 624  
 : TYPE: PRT

```

: ORGANISM: Homo sapiens
US-10-002-050-22
Query Match      96.5%; Score 3107.5; DB 15; Length 624;
Best Local Similarity 99.0%; Pred. No. 6.1e-307;
Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 MAPHMAVMLLAARLKMGLGIGAEYWMNLVPRKTVSSSELTAVRRFSQGTGIDELTTLTE 60
DB 1 MAPHMAVMLLAARLKMGLGIGAEYWMNLVPRKTVSSSELTAVRRFSQGTGIDELTTLTE 60
OY 61 PTGLLVGAREALFAFASMEALDELQAIISWAPAEKTECIQKGNKNOTECFNIRLQAP 120
DB 61 PTGLLVGAREALFAFASMEALDELQAIISWAPAEKTECIQKGNKNOTECFNIRLQAP 120
OY 121 NASHLYVCGTYAFAPQCTYVNMFLTLEHGEFEDGKCKPYDPKAKHAGLLVDELXSA 180
DB 121 NASHLYVCGTYAFAPQCTYVNMFLTLEHGEFEDGKCKPYDPKAKHAGLLVDELXSA 180
OY 181 LNPFLETEPIILNMGPHHSKTEYLAFWLNBPHEFGASAVPESVGSFTGDDDKVYFFFR 240
DB 181 LNPFLETEPIILNMGPHHSKTEYLAFWLNBPHEFGASAVPESVGSFTGDDDKVYFFFR 240
OY 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
DB 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
OY 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVEGPRYKEYHEAOKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVEGPRYKEYHEAOKMDRYTD 360
OY 361 PVSPRRGSCINNMHRRHGTSSLELPDNLNPKVKKHPLMEBOVGRRMSRPLLYKKGTFN 420
DB 361 PVSPRRGSCINNMHRRHGTSSLELPDNLNPKVKKHPLMEBOVGRRMSRPLLYKKGTFN 420
OY 421 THLYADRYTGLDGAATYTLVFIGTGDGMLKAVSLGPMVHLIEELQLEFQDPKMSLVLS 480
DB 421 THLYADRYTGLDGAATYTLVFIGTGDGMLKAVSLGPMVHLIEELQLEFQDPKMSLVLS 480
OY 481 KILLFAGSSQVQLPVADCIKRYSCADCVLARDPYCAMSVMVSRCAVAGHGSLILQH 540
DB 481 KILLFAGSSQVQLPVADCIKRYSCADCVLARDPYCAMSVMVSRCAVAGHGSLILQH 540
OY 541 VMTSDTSGICNLGRSKIQSGPRKNTTVAGTDVLPCHLSS 582
DB 541 VMTSDTSGICNLGRSK-KVPRPKNTTVAGTDVLPCHLSS 581

RESULT 10
US-10-002-304-22
: Sequence 22, Application US/10002304
: Publication No. US20030036185A1
: GENERAL INFORMATION:
: APPLICANT: Shinkels, Richard
: APPLICANT: Fernandes, Elma
: APPLICANT: Vernet, Corine
: APPLICANT: Yang, Meljia
: APPLICANT: Boldog, Ferenc
: APPLICANT: Heitmann, John
: TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
: FILE REFERENCE: 15966-554 Cura-54 CON-S8
: CURRENT APPLICATION NUMBER: US/10/002,304
: PRIORITY FILING DATE: 2001-11-02
: PRIOR APPLICATION NUMBER: 09/604,286
: PRIOR FILING DATE: 2000-06-22
: PRIOR APPLICATION NUMBER: 60/140,584
: PRIOR FILING DATE: 1999-06-23
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 22
: LENGTH: 624
: TYPE: PRF
: ORGANISM: Homo sapiens
```

```

US-10-002-304-22
Query Match      96.5%; Score 3107.5; DB 15; Length 624;
Best Local Similarity 99.0%; Pred. No. 6.1e-307;
Matches 576; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 MAPHMAVMLLAARLKMGLGIGAEYWMNLVPRKTVSSSELTAVRRFSQGTGIDELTTLTE 60
DB 1 MAPHMAVMLLAARLKMGLGIGAEYWMNLVPRKTVSSSELTAVRRFSQGTGIDELTTLTE 60
OY 61 PTGLLVGAREALFAFASMEALDELQAIISWAPAEKTECIQKGNKNOTECFNIRLQAP 120
DB 61 PTGLLVGAREALFAFASMEALDELQAIISWAPAEKTECIQKGNKNOTECFNIRLQAP 120
OY 121 NASHLYVCGTYAFAPQCTYVNMFLTLEHGEFEDGKCKPYDPKAKHAGLLVDELXSA 180
DB 121 NASHLYVCGTYAFAPQCTYVNMFLTLEHGEFEDGKCKPYDPKAKHAGLLVDELXSA 180
OY 181 LNPFLETEPIILNMGPHHSKTEYLAFWLNBPHEFGASAVPESVGSFTGDDDKVYFFFR 240
DB 181 LNPFLETEPIILNMGPHHSKTEYLAFWLNBPHEFGASAVPESVGSFTGDDDKVYFFFR 240
OY 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
DB 241 ERAVESDCYAEQVAVARVCKDGMGARTLQKRTTFLKARLACGAPNMQLYFNOLQAM 300
OY 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVEGPRYKEYHEAOKMDRYTD 360
DB 301 HTLQDTSMHNTTFEGVFOAQMGDMYLSAICEYOLEIQRVEGPRYKEYHEAOKMDRYTD 360
OY 361 PVSPRRGSCINNMHRRHGTSSLELPDNLNPKVKKHPLMEBOVGRRMSRPLLYKKGTFN 420
DB 361 PVSPRRGSCINNMHRRHGTSSLELPDNLNPKVKKHPLMEBOVGRRMSRPLLYKKGTFN 420
OY 421 THLYADRYTGLDGAATYTLVFIGTGDGMLKAVSLGPMVHLIEELQLEFQDPKMSLVLS 480
DB 421 THLYADRYTGLDGAATYTLVFIGTGDGMLKAVSLGPMVHLIEELQLEFQDPKMSLVLS 480
OY 481 KILLFAGSSQVQLPVADCIKRYSCADCVLARDPYCAMSVMVSRCAVAGHGSLILQH 540
DB 481 KILLFAGSSQVQLPVADCIKRYSCADCVLARDPYCAMSVMVSRCAVAGHGSLILQH 540
OY 541 VMTSDTSGICNLGRSKIQSGPRKNTTVAGTDVLPCHLSS 582
DB 541 VMTSDTSGICNLGRSK-KVPRPKNTTVAGTDVLPCHLSS 581

RESULT 11
US-09-759-130B-335
: Sequence 335, Application US/09759130B
: Publication No. US20030022279A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: APPLICANT: McCarthy, Sean A
: APPLICANT: Fraser, Christopher C
: APPLICANT: Sharp, John D
: APPLICANT: Barnes, Thomas S
: APPLICANT: Kirst, Susan J
: APPLICANT: Mackay, Charles R
: APPLICANT: Leiby, Kevin R
: APPLICANT: Wrighton, Nicolas
: APPLICANT: Goodgar, Andrew
: APPLICANT: Holtzman, Douglas A
: TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
: FILE REFERENCE: MP100-3550NM1M
: CURRENT APPLICATION NUMBER: US/09/759,130B
: CURRENT FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: US 09/479,249
: PRIOR FILING DATE: 2000-01-07
: PRIOR APPLICATION NUMBER: US 09/559,497
```



```

: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 09/578,063
: PRIOR FILING DATE: 2000-05-24
: PRIOR APPLICATION NUMBER: US 09/333,159
: PRIOR FILING DATE: 1999-06-14
: PRIOR APPLICATION NUMBER: US 09/596,194
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 09/342,364
: PRIOR FILING DATE: 1999-06-29
: PRIOR APPLICATION NUMBER: US 09/608,452
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/393,996
: PRIOR FILING DATE: 1999-09-10
: PRIOR APPLICATION NUMBER: US 09/602,871
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 09/420,707
: PRIOR FILING DATE: 1999-10-19
: NUMBER OF SEQ ID NOS: 460
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 335
: LENGTH: 834
: TYPE: PRT
: ORGANISM: Mus sp.
US-09-759-1308-335

```

```

Query Match      84.5%; Score 2720; DB 11; Length 834;
Best Local Similarity 83.1%; Pred. No. 2.5e-267;
Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;

```

```

QY 1 MAPHMAVWLLAARLGMGIGAGAEVWNNLVPRKTVSSGELATVYRRFSQTGIDFILTTLTE 60
DB 1 MAPHMAVWLLAAGLGMGIGAGAEVWNNLVPRKTVSSGELATVYRRFSQTGIDFILTTLTE 60
QY 61 PTGLLYVGAREALPAFSEALELOGAISWEAPVEKTECIQKGNNOECFNFTRFLQPY 120
DB 61 HSGLLYGAREALPAFSEALELOGAISWEAPVEKTECIQKGNNOECFNFTRFLQPY 120
QY 121 NASHLVYCGTYAFQPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAGLLVDELVSAT 180
DB 121 NSHLVYCGTYAFQPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAGLLVDELVSAT 180
QY 181 LNNTLGEPIILRMNGHSHKTEYLAFWLNPHFVSAYPESVGSFTGDDKVFYFFR 240
DB 181 LNNTLGEPIILRMNGHSHKTEYLAFWLNPHFVSAYPESVGSFTGDDKVFYFFR 240
QY 241 ERAVESCYAEQVYARAVYCKGDMGARTLQKRTTFLKARLACSAAPNQLYFNOLQAM 300
DB 241 ERAVESCYAEQVYARAVYCKGDMGARTLQKRTTFLKARLACSAAPNQLYFNOLQAM 300
QY 301 HTLQDTSNHTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQYKEYHEBAQKMDRYTD 360
DB 301 HTLQDTSNHTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQYKEYHEBAQKMDRYTD 360
QY 361 PVSPRPGSCINNMHRHNGYSSLELDPNLTINFAKRPILMEQVGPMSRPLVKKGNF 420
DB 361 PVSPRPGSCINNMHRHNGYSSLELDPNLTINFAKRPILMEQVGPMSRPLVKKGNF 420
QY 421 THLVADRYTGLDGTATYVLFITGDMGLKAVSLGPWHLIEELDLFQDEMRSLVLSQS 480
DB 421 THLVADRYTGLDGTATYVLFITGDMGLKAVSLGPWHLIEELDLFQDEMRSLVLSQS 480
QY 481 KKLIFAGRSQDLVOLPVADCIKYSACDVLARPYCAMSVNTSRCAV-VGHHGSLILQ 539
DB 481 KKLIFAGRSQDLVOLPVADCIKYSACDVLARPYCAMSVNTSRCAV-VGHHGSLILQ 539
QY 540 HVMTSDFSGICNLNG-SKIOSGPXPKNTTVAGTDLVLPCHLSSNLA-----LPD 588
DB 540 HVMTSDFSGICNLNG-SKIOSGPXPKNTTVAGTDLVLPCHLSSNLA-----LPD 588
QY 541 HVANLDTSKMCKNOYGIKKVRS--IPKNTTVSGTDLVLPCHLSSNLAHMTFGSODLPA 598
DB 541 HVANLDTSKMCKNOYGIKKVRS--IPKNTTVSGTDLVLPCHLSSNLAHMTFGSODLPA 598
QY 589 SNP 591
DB 589 EOP 601

```

```

RESULT 12
US-10-189-123-65
: Sequence 65, Application US/10189123
: Publication NO. US20030082586A1
: GENERAL INFORMATION:
: APPLICANT: KIRST, Susan J.
: APPLICANT: HOLTZMAN, Douglas A.
: APPLICANT: FRASER, Christopher C.
: APPLICANT: SHARP, John D.
: APPLICANT: BARNES, Thomas S.
: TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
: FILE REFERENCE: 10147-1103
: CURRENT APPLICATION NUMBER: US/10/189,123
: CURRENT FILING DATE: 2002-07-02
: PRIOR APPLICATION NUMBER: US 09/596,194
: PRIOR FILING DATE: 2000-06-16
: PRIOR APPLICATION NUMBER: US 09/342,364
: PRIOR FILING DATE: 1999-06-29
: NUMBER OF SEQ ID NOS: 100
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 65
: LENGTH: 834
: TYPE: PRT
: ORGANISM: Mus sp.
US-10-189-123-65

```

```

Query Match      84.5%; Score 2720; DB 15; Length 834;
Best Local Similarity 83.1%; Pred. No. 2.5e-267;
Matches 501; Conservative 41; Mismatches 47; Indels 14; Gaps 4;

```

```

QY 1 MAPHMAVWLLAARLGMGIGAGAEVWNNLVPRKTVSSGELATVYRRFSQTGIDFILTTLTE 60
DB 1 MAPHMAVWLLAAGLGMGIGAGAEVWNNLVPRKTVSSGELATVYRRFSQTGIDFILTTLTE 60
QY 61 PTGLLYVGAREALPAFSEALELOGAISWEAPVEKTECIQKGNNOECFNFTRFLQPY 120
DB 61 HSGLLYGAREALPAFSEALELOGAISWEAPVEKTECIQKGNNOECFNFTRFLQPY 120
QY 121 NASHLVYCGTYAFQPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAGLLVDELVSAT 180
DB 121 NSHLVYCGTYAFQPKCTYVNMFLTELEHGEFEDGKGCPCYDPAKGAGLLVDELVSAT 180
QY 181 LNNTLGEPIILRMNGHSHKTEYLAFWLNPHFVSAYPESVGSFTGDDKVFYFFR 240
DB 181 LNNTLGEPIILRMNGHSHKTEYLAFWLNPHFVSAYPESVGSFTGDDKVFYFFR 240
QY 241 ERAVESCYAEQVYARAVYCKGDMGARTLQKRTTFLKARLACSAAPNQLYFNOLQAM 300
DB 241 ERAVESCYAEQVYARAVYCKGDMGARTLQKRTTFLKARLACSAAPNQLYFNOLQAM 300
QY 301 HTLQDTSNHTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQYKEYHEBAQKMDRYTD 360
DB 301 HTLQDTSNHTTFFGVFOAQMGDMYLSAICEYOLEIQRVFEQYKEYHEBAQKMDRYTD 360
QY 361 PVSPRPGSCINNMHRHNGYSSLELDPNLTINFAKRPILMEQVGPMSRPLVKKGNF 420
DB 361 PVSPRPGSCINNMHRHNGYSSLELDPNLTINFAKRPILMEQVGPMSRPLVKKGNF 420
QY 421 THLVADRYTGLDGTATYVLFITGDMGLKAVSLGPWHLIEELDLFQDEMRSLVLSQS 480
DB 421 THLVADRYTGLDGTATYVLFITGDMGLKAVSLGPWHLIEELDLFQDEMRSLVLSQS 480
QY 481 KKLIFAGRSQDLVOLPVADCIKYSACDVLARPYCAMSVNTSRCAV-VGHHGSLILQ 539
DB 481 KKLIFAGRSQDLVOLPVADCIKYSACDVLARPYCAMSVNTSRCAV-VGHHGSLILQ 539
QY 540 HVMTSDFSGICNLNG-SKIOSGPXPKNTTVAGTDLVLPCHLSSNLA-----LPD 588
DB 540 HVMTSDFSGICNLNG-SKIOSGPXPKNTTVAGTDLVLPCHLSSNLA-----LPD 588
QY 541 HVANLDTSKMCKNOYGIKKVRS--IPKNTTVSGTDLVLPCHLSSNLAHMTFGSODLPA 598
DB 541 HVANLDTSKMCKNOYGIKKVRS--IPKNTTVSGTDLVLPCHLSSNLAHMTFGSODLPA 598
QY 589 SNP 591
DB 589 EOP 601

```

Db 599 EOP 601

RESULT 13

US-09-989-920-218

Sequence 218, Application US/09989920

Patent No. US2002012957A1

GENERAL INFORMATION:

APPLICANT: Macina, Roberto

APPLICANT: Recipon, Herve

APPLICANT: Chen, Sai-Yu

APPLICANT: Sun, Yongming

APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro

FILE REFERENCE: DEX-0291

CURRENT APPLICATION NUMBER: US/09/989,920

CURRENT FILING DATE: 2001-11-21

PRIOR APPLICATION NUMBER: 60/252,500

PRIOR FILING DATE: 2000-11-22

NUMBER OF SEQ ID NOS: 284

SOFTWARE: PatentIn version 3.1

SEQ ID NO 218

LENGTH: 785

TYPE: PRT

ORGANISM: Homo sapien

US-09-989-920-218

Query Match 81.9%; Score 2635; DB 10; Length 785;

Best Local Similarity 90.0%; Pred. No. 1e-258;

Matches 495; Conservative 2; Mismatches 7; Indels 46; Gaps 4;

QY 87 ISWEARVEKTECICOKKNN-----OQECN 112

DB 4 ISWEARVEKTECICOKKNNQVANTLLVLPSPDVSSHSPRALNTRPCPOTECH 63

QY 113 FIRLPDPYASHLYVCGTAFOPKCTYVNLFTLEHGEFEDEKGCSPDPKAGHGLY 172

DB 64 FIRLPDPYASHLYVCGTAFOPKCTYVNLFTLEHGEFEDEKGCSPDPKAGHGLY 123

QY 173 DGLYSATLNLGTEPILRNKGRHNSKTEYLFWLNEPFFVGSAYVESVSEFTGDD 232

DB 124 DEELYSATLNLGTEPILRNKGRHNSKTEYLFWLNEPFFVGSAYVESVSEFTGDD 183

QY 233 DAVYFFRRRAVESDCYAOVYARAVARVCKGDMGARTLQKRWTFPLKRLACSAFNMQL 292

DB 184 DAVYFFRRRAVESDCYAOVYARAVARVCKGDMGARTLQKRWTFPLKRLACSAFNMQL 243

QY 293 YFNQLQAMHTLQDTSWNTTFFGVFOAGMDYLSAICEYOLELQORVEGPKYKEHNEA 352

DB 244 YFNQLQAMHTLQDTSWNTTFFGVFOAGMDYLSAICEYOLELQORVEGPKYKEHNEA 303

QY 353 QKMDRYTDPVSPRPGSCINNNHRRHGYTSSLELDPNLTNFYKKHPLMEEOVGPRMSRPL 412

DB 304 QKMDRYTDPVSPRPGSCINNNHRRHGYTSSLELDPNLTNFYKKHPLMEEOVGPRMSRPL 363

QY 413 LVKKGSTNTHLVADAVTGLDGAATYVLFCTGDDGMLKRAVSLGPRVHLLLEELQLDQDEM 472

DB 364 LVKKGSTNTHLVADAVTGLDGAATYVLFCTGDDGMLKRAVSLGPRVHLLLEELQLDQDEM 423

QY 473 RSLVYSQSK-KLFFAGSRSLVQLVPAACIKRSCADCVLADPVCAMSVNTRCAVAGC 531

DB 424 RSLVYSQSKVLLFFAGSRSLVQLVPAACIKRSCADCVLADPVCAMSVNTRCAVAGC 483

QY 532 HEGSLLIHWMTSDTSGICNLGSKISQSPKPKNTTVAGTDVLPCHLSSNLA----- 585

DB 484 HEGSLLIHWMTSDTSGICNLGSK-KVAPPTPKNTTVAGTDVLPCHLSSNLAHARWTF 542

QY 586 ----LPDSNP 591

DB 543 GGRDLPAEOP 552

RESULT 14

US-09-946-374-253

Sequence 253, Application US/09946374

Publication No. US20030073129A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Pan, James

APPLICANT: Proul, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2830P1C1

CURRENT APPLICATION NUMBER: US/09/946,374

CURRENT FILING DATE: 2001-09-04

PRIOR APPLICATION NUMBER: 60/098716

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098723

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098749

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098750

PRIOR FILING DATE: 1998-09-01

PRIOR APPLICATION NUMBER: 60/098803

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/098821

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/098843

PRIOR FILING DATE: 1998-09-02

PRIOR APPLICATION NUMBER: 60/099536

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099596

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099598

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099602

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099642

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099741

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099754

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099763

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099792

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099808

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099812

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099815

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099816

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/100385

PRIOR FILING DATE: 1998-09-15

PRIOR APPLICATION NUMBER: 60/100388

;; PRIOR FILING DATE: 1998-09-15  
;; PRIOR APPLICATION NUMBER: 60/100390  
;; PRIOR FILING DATE: 1998-09-15  
;; PRIOR APPLICATION NUMBER: 60/100584  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: 60/100627  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: 60/100661  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: 60/100662  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: 60/100664  
;; PRIOR FILING DATE: 1998-09-16  
;; PRIOR APPLICATION NUMBER: 60/100683  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/100684  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/100710  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/100711  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/100848  
;; PRIOR FILING DATE: 1998-09-18  
;; PRIOR APPLICATION NUMBER: 60/100849  
;; PRIOR FILING DATE: 1998-09-18  
;; PRIOR APPLICATION NUMBER: 60/100919  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/100930  
;; PRIOR FILING DATE: 1998-09-17  
;; PRIOR APPLICATION NUMBER: 60/101014  
;; PRIOR FILING DATE: 1998-09-18  
;; PRIOR APPLICATION NUMBER: 60/101068  
;; PRIOR FILING DATE: 1998-09-18  
;; PRIOR APPLICATION NUMBER: 60/101071  
;; PRIOR FILING DATE: 1998-09-18  
;; PRIOR APPLICATION NUMBER: 60/101279  
;; PRIOR FILING DATE: 1998-09-22  
;; PRIOR APPLICATION NUMBER: 60/101471  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101472  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101474  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101475  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101476  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101477  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101479  
;; PRIOR FILING DATE: 1998-09-23  
;; PRIOR APPLICATION NUMBER: 60/101738  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/101741  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/101743  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/101915  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/101916  
;; PRIOR FILING DATE: 1998-09-24  
;; PRIOR APPLICATION NUMBER: 60/102207  
;; PRIOR FILING DATE: 1998-09-29  
;; PRIOR APPLICATION NUMBER: 60/102240  
;; PRIOR FILING DATE: 1998-09-29  
;; PRIOR APPLICATION NUMBER: 60/102307  
;; PRIOR FILING DATE: 1998-09-29  
;; PRIOR APPLICATION NUMBER: 60/102330  
;; PRIOR FILING DATE: 1998-09-29  
;; PRIOR APPLICATION NUMBER: 60/102331  
;; PRIOR FILING DATE: 1998-09-29  
;; PRIOR APPLICATION NUMBER: 60/102484  
;; PRIOR FILING DATE: 1998-09-30

;; PRIOR APPLICATION NUMBER: 60/102487  
;; PRIOR FILING DATE: 1998-09-30  
;; PRIOR APPLICATION NUMBER: 60/102570  
;; PRIOR FILING DATE: 1998-09-30  
;; PRIOR APPLICATION NUMBER: 60/102571  
;; PRIOR FILING DATE: 1998-09-30  
;; PRIOR APPLICATION NUMBER: 60/102684  
;; PRIOR FILING DATE: 1998-10-01  
;; PRIOR APPLICATION NUMBER: 60/102687  
;; PRIOR FILING DATE: 1998-10-01  
;; PRIOR APPLICATION NUMBER: 60/102965  
;; PRIOR FILING DATE: 1998-10-02  
;; PRIOR APPLICATION NUMBER: 60/103258  
;; PRIOR FILING DATE: 1998-10-06  
;; PRIOR APPLICATION NUMBER: 60/103314  
;; PRIOR FILING DATE: 1998-10-07  
;; PRIOR APPLICATION NUMBER: 60/103315  
;; PRIOR FILING DATE: 1998-10-07  
;; PRIOR APPLICATION NUMBER: 60/103328  
;; PRIOR FILING DATE: 1998-10-07  
;; PRIOR APPLICATION NUMBER: 60/103395  
;; PRIOR FILING DATE: 1998-10-07  
;; PRIOR APPLICATION NUMBER: 60/103396  
;; PRIOR FILING DATE: 1998-10-07  
;; PRIOR APPLICATION NUMBER: 60/103401  
;; PRIOR FILING DATE: 1998-10-07  
;; PRIOR APPLICATION NUMBER: 60/103449  
;; PRIOR FILING DATE: 1998-10-06  
;; PRIOR APPLICATION NUMBER: 60/103633  
;; PRIOR FILING DATE: 1998-10-08  
;; PRIOR APPLICATION NUMBER: 60/103678  
;; PRIOR FILING DATE: 1998-10-08  
;; PRIOR APPLICATION NUMBER: 60/103679  
;; PRIOR FILING DATE: 1998-10-08  
;; PRIOR APPLICATION NUMBER: 60/103711  
;; PRIOR FILING DATE: 1998-10-08  
;; PRIOR APPLICATION NUMBER: 60/104257  
;; PRIOR FILING DATE: 1998-10-14  
;; PRIOR APPLICATION NUMBER: 60/104987  
;; PRIOR FILING DATE: 1998-10-20  
;; PRIOR APPLICATION NUMBER: 60/105000  
;; PRIOR FILING DATE: 1998-10-20  
;; PRIOR APPLICATION NUMBER: 60/105002  
;; PRIOR FILING DATE: 1998-10-20  
;; PRIOR APPLICATION NUMBER: 60/105104  
;; PRIOR FILING DATE: 1998-10-21  
;; PRIOR APPLICATION NUMBER: 60/105169  
;; PRIOR FILING DATE: 1998-10-22  
;; PRIOR APPLICATION NUMBER: 60/105266  
;; PRIOR FILING DATE: 1998-10-22  
;; PRIOR APPLICATION NUMBER: 60/105693  
;; PRIOR FILING DATE: 1998-10-26  
;; PRIOR APPLICATION NUMBER: 60/105694  
;; PRIOR FILING DATE: 1998-10-26  
;; PRIOR APPLICATION NUMBER: 60/105807

Query Match 38.5% Score 1240: DB 11: Length 837:  
Best Local Similarity 45.4% Pred. No 8.8e-117:  
Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12:

OY 25 WNLVPRKTVSSGELATVVRPSQGTIGDPLTTLTEPTGLLYVGAREALPAFMSKALRIQ 84  
DB 42 WALSPRISLPLGSEERPLRFEAEHISVYTAALLSROGRTLYVGAREALPLSSLSLPL 101  
OY 85 GA-----ISWEAPVEKKTICIOKGNNOCTECNFTRFLOPYANASHLYVCGYTAPOPKTYV 140  
DB 102 GGEYOELLMGADARKKQCSFKDPODQCONYIKILPLSGSHLFTCTGTAFFSMCTYI 161  
OY 141 NMLFTLEHGE-----FEDGKGCPCYPDAKGHAGLIVGELYSATINFLGTEPIILNM 195  
DB 162 MNEFTTLARDKGNVLEDDGKRCRPFDPNFTSTALVYDGEIYGTGVSSFGQNDPAISNSQ 221  
OY 196 GPHSMKTEYLAFMLNEPHFVSAVYVESVSGFTGDDDKVYFFFRERAVESDCYAQVYA 255

```
Db 222 SLRPT KTSSLMWLODPFVAVASATIPESLGSLOQDDDKTYFFSFTGOEFFEENTIVS 280
OY 256 RVARVCGDMGARGATLOKRTTEFLKARLACSAFNMOLYFNQLOAMTLODT--SMNTEF 313
Db 281 RIARICKGDEGGRVLOQWHTSFLKQLOLCSRPDDFFPNVLODVTLSSPDMDITLF 340
OY 314 FGVCQAW--GDMYLSAICEYOLEIORYPECKYKHEBAOKMDRYTDPVSPRSCSI 371
Db 341 YGVFTSQWHRGTTEGSAVCVFTMKDVQAFVSGLYKKEVNETOQMYTVTRPVPTRGACI 400
OY 372 NNMHRBGYTSLSLELDNLIINFVKKHLMBEOYGRPMSPILYKKGTFNTHLVADVTGL 431
Db 401 TMSARBRKINSSLOLDRVNLKDFLMDGOVR--SMMLLOQPAROYRAVHRVPGI 457
OY 432 DGATYVLEIGTGDGMLKAVSLGPMVHLEIQLFDO--EPMSLVLSQSKLLFAGSRS 490
Db 458 H-HTYDVFLGTGDXLHKAVSIGPRVHIELOITSSQOPQNLDDHRLGLLYAHSIS 516
OY 491 QLVQPLVADCIKTRSCADCVLARDPYCANSVNTSRCAVANGHFGSL--IOHVTST 546
Db 517 GVOYVMAVCSLYRSGDCLLARDPYCANS--GSSCKHVSLOPOLATRPWIODIGASA 574
OY 547 SGICN---LRGSKIQSGXP--KNITVAGTDLVLPCHLSNLA 585
Db 575 KDLCSASSVSPSFPVTGKPCBOVOPOPNTVNTLACPLISNLA 618

RESULT 15
US-10-052-586-454
; Sequence 454, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William J.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31

; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
```

```

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12

```

```

; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

```

```

Query Match 38.5%; Score 1240; DB 14; Length 837;
Best Local Similarity 45.4%; Pred. No. 8.8e-117;
Matches 265; Conservative 88; Mismatches 201; Indels 30; Gaps 12;

```

```

QY 25 WNLVPRKTVSSGELATVYRRSQCIGDPLTLTEPLGILLYGAREALFSPSEALQL 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 42 WALSPRSLPGRSEPRPLREAEHISNYTALLSRDGRILYVGAREALFALSSNLSFLP 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 85 GA---ISMEAPVEKTECIQKNNOTECNFTFPIQPYNAHLYVGTAFQPKCYV 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102 GGEYQELMGADAEKKQCSFKGDPQDCONTIKILPLSGSHLFTCGTAFSPMCTYI 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 NMLFTLEHGE---FEDGKGCPRYPKAGHAGLYVDELYSATLNNFTLGTETILRNM 195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 162 NMENFTLARDEKGNVLTEDGRCPDPENFSTALYVDELITGTVSSFGQNDPAISRSQ 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 GPHSMKTEYLAFNINEHPVGSAYVPSVSGSTGGDDKTYFFFRERAVESDCYAEQVYA 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 222 SLRPT-KTESSLNMLDPAFVAVASVIPSLSLQDDDKIYFFSEFGQEFEPENTIVS 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 RVARCKGDMGCAITLQKRTTFLKARLACGAPNMQLYFNOLQAMHTLQDT--SMHNTF 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 281 RIARCKDBEGGERLQGRMTSFLKAOLCSRPDGFPFNVLQDYFTLSPSPQDRDRLTF 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 314 FGVPQAOV--GDMTSLAICEYQLEIQRVEFGPYKEVHEAOKMDRYTDPVSPRPGSCI 371
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 341 YGVEFSQWHRGTTEGSAVCFVTKMDQVRFSGLYKEVARETQWYTVTHPVPTPRGACI 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 372 NMHRHGHYTSLSLEPDLILFVYKHPLEMGQVGRMRPLLYKKGTFTLADRYTGL 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 401 TNSAERRKINSLLQPDVYVNFLLKDHFLMDQVYR--SRMLLLQPARYQVAVARVYGL 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 432 DCAATYVLFITGDMWLKAVSLGPMVHLEIQLDFD--EPMRSVLVSOSKLLFAGRS 490
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 458 H-HYDVLFLATGDCRLKAVSVGPRVHIEELQIFSSGQVQVONLLDTHGGLTAASHS 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 491 QLVOLPVADCIKYSACDCVILADPYCAMSVNTSKCVAVGGHFGSLT---IOHMYTSDT 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 517 GVOQVPMANGSLYRSCGCLLARDPYCAMS--GSSCKHYSLYQPLATRPWIDIEGASA 574
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 SGICH---LNGSKIQSGPXP-KNITVYVAGDVLVPCHLSSNYLA 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 575 KDLCSASSVSPSPFVTEKPECEVOQOPNTVNTLACPLLSNTLA 618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Search completed: August 7, 2003, 13:59:11
Job time : 56 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 13:13:42 ; Search time 3207 Seconds

(without alignments)  
17309.455 Million cell updates/sec

Title: us-10-002-050-13

Perfect score: 2284  
Sequence: 1 cggcccttcacacactctg.....tgtgtgtgtgtatgtgtgtg 2284

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estda:\*  
2: em\_estnum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rnd:\*  
26: em\_gss\_ping:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303.8	57.1	3648	AK090073	Mus muscu
2	832.4	36.4	977	BX370082	BX370082
3	831.8	36.4	932	BX369159	BX369159
4	688	30.1	864	B1518922	B1518922

5	651.4	28.5	929	13	BQ917819	BQ917819
6	632.6	27.7	705	10	BE294211	BE294211
7	621.2	27.2	937	13	BQ154848	BQ154848
8	603.8	26.4	1394	12	BM926974	BM926974
9	597.6	26.2	768	12	BM943605	BM943605
10	597	26.1	632	10	BE270240	BE270240
11	579.6	25.4	1002	13	BQ068313	BQ068313
12	575.4	25.2	853	12	BX396978	BX396978
13	569.2	24.9	745	13	BM626964	BM626964
14	554.6	24.3	763	12	BQ161234	BQ161234
15	536.4	23.5	597	10	BQ324157	BQ324157
16	536	23.5	1075	13	BX360753	BX360753
17	527.4	23.1	978	13	BQ963608	BQ963608
18	520.2	22.8	738	12	BM414643	BM414643
19	507.4	22.2	965	13	BQ845252	BQ845252
20	501.8	22.0	977	13	BQ855862	BQ855862
21	494.8	21.7	647	12	BM950623	BM950623
22	487.6	21.3	673	12	BM951984	BM951984
23	477	20.9	789	12	B1686539	B1686539
24	452	19.8	1033	13	BX460723	BX460723
25	443.4	19.4	604	10	BE283742	BE283742
26	435.4	19.1	455	14	W01259	W01259
27	435	19.0	1129	13	BQ072179	BQ072179
28	431.2	18.9	766	12	B1643497	B1643497
29	428	18.7	533	13	BQ266902	BQ266902
30	428	18.7	923	12	B165624	B165624
31	397.8	17.4	704	14	BY741417	BY741417
32	395.4	17.3	710	9	A1651881	A1651881
33	391.2	17.1	840	14	CB990331	CB990331
34	390.4	17.1	404	12	BM174889	BM174889
35	387.6	17.0	911	13	BQ278955	BQ278955
36	378.6	16.6	3763	11	AK088653	AK088653
37	376.8	16.5	2777	11	AK028896	AK028896
38	376.2	16.5	659	9	A1651838	A1651838
39	375.2	16.4	3809	11	AK030331	AK030331
40	372.4	16.3	707	14	BY733150	BY733150
41	367	16.1	690	14	BY734086	BY734086
42	358.8	15.7	573	10	BE070335	BE070335
43	343.4	15.0	453	14	CB060019	CB060019
44	340	14.9	680	14	BY735266	BY735266
45	336.2	14.7	543	14	CD216128	CD216128

#### ALIGNMENTS

-RESULT 1  
AK090073  
LOCUS  
DEFINITION  
Mus musculus 7 days embryo nullipotent stem cell CRL-2070 NE CDNA,  
RIKEN full-length enriched library, clone:G430091K07 product:sema  
domain, immunoglobulin domain (Ig), transmembrane domain (TM) and  
short cytoplasmic domain, (semaphorin) 4C, full insert sequence.  
ACCESSION  
AK090073  
VERSION  
AK090073.1 GI:26105718  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
Carninci,P. and Hayashizaki,Y.  
TITLE  
High-efficiency full-length cDNA cloning  
METH. ENZYMOL. 303, 19-44 (1999)  
MEDLINE  
99279253  
PUBMED  
10349636  
REFERENCE  
AUTHORS  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL  
Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE  
20499374

PUBMED  
REFERENCE  
AUTHORS  
11042159  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Kono,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujikawa,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,  
Tameda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED  
11076861

REFERENCE  
AUTHORS  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
6 (bases 1 to 3648)  
Adachi,J., Aizawa,K., Akimura,T., Arikawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M.,  
Hayashida,K., Hayata,N., Hiramoto,K., Hirooka,T., Hirozane,T.,  
Horii,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M.,  
Koye,S., Kurimura,C., Matsuyama,T., Miyazaki,A., Murata,H.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akihira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsukumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/.

FEATURES  
SOURCE  
Location/Qualifiers  
1..3648  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C3H/J"  
/db\_xref="PANTOM,DB:G430091K07"  
/db\_xref="taxon:10090"  
/clone="G430091K07"  
/cell\_line="CRL-2070 NE"  
/cell\_type="nullipotent stem cell"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="7 days embryo"  
245..2752  
/note="putative  
sema domain, Immunoglobulin domain (19), transmembrane  
domain (TM) and short cytoplasmic domain, (semaphorin) 4C  
(MGI:109252, GBS159463, evidence: BLASTN, 98%,  
match=3432)."

BASE COUNT  
728 a 985 c 1089 g 844 t 2 others

ORIGIN  
Query Match 57.1%; Score 1303.8; DB 11; Length 3648;  
Best Local Similarity 83.4%; Pred. No. 3,2e+290;  
Matches 1542; Conservative 0; Mismatches 298; Indels 10; Gaps 5;  
QY 97 GGACCTCCGCTCGCAGCCGGCCGACAGAGACAGTGGCTGAGCTCAGAGCCGGG 156  
DB 176 GCACACACCGGGGTTTGTAGCCCCAGCTTAGGACAGGGGTGCTGAGCTGAGATTGAGACCTGG 235  
QY 157 GCGTCGGCCATGGCCCCACACAGCTGGCTGTGGCTGGGACAGCAAGCTGGGGCCCTG 216  
DB 236 GTGTGCTGACAGTGGCCCCACAGCTGGCTGTGGCTGGGACAGCAAGCTGGGGCCCTG 295  
QY 217 GGCATTGGGGCTGAGGTGTGTGAGACCTGTGCGCCGTAGACAGATGCTTCTTGGGAG 276  
DB 296 GGCATCGGGGGCTGAGATGTGTGAGACCTTGTGGCCCGGAGACAGATCTTCTGGGAG 355  
QY 277 CTGGCCACGAGTATGAGCGGGTCTCCAGACAGCCGATCCAGAGACTTCTGACACTGACG 336  
DB 356 CTGGTCACAGTATGAGCGGGTCTCCAGACAGCCGATCCAGAGACTTCTGACACTGACG 415  
QY 337 CTGACGGAGCCCACTGGGGCTTCTGACGTGGGGCCCGAGAGCCCTGTTGGCTTACG 396  
DB 416 CTGACAGAACATTTGGCTTTATATGTGTGGGGCCGAGAGGCGCTTGTGCTTCACT 475  
QY 397 ATGAGGCCCTGAGACTGCAAGAGAGATCTCTGAGAGCCCGCTGGAGAGAGACT 456  
DB 476 GTAGAGGCTCTGAGACTGCAAGAGAGATCTCTGAGAGCCCGCTGGAGAGAGATTT 535  
QY 457 GAGTGTATCCAGAAAGGAGAGAGACAGACAGAGCCGATCTTCACTTCACTCCGCTCTG 516  
DB 536 GAAATATACCGAAAGGAGAGAGACAGACAGAGCCGATCTTCACTTCACTCCGCTCTG 595  
QY 517 CAGCCCTCAAA-TGGCTCCCACTCTGAGCTGTGGACAGCTTCCAGCCCACTTGG 575  
DB 596 CAGCCATTAACCTTCTCTCCATCTGTATGTGTGCGGCACTTATGCTTCCAGCCCACTG 655  
QY 576 CACCTAGCTCAACATGCTACCTTCTGAGCATGAGAGTGTGAAGATGGAGAGG 635  
DB 656 CACCTAGCTCAACATGCTACCTTCTGAGCATGAGAGTGTGAAGATGGAGAGG 715  
QY 636 CAGTGTCTCTATGACCCAGCTAAGAGGCTTGTGGCTTCTTGTGATGTGACCTGTA 695  
DB 716 TAAATGCCCAATATGACCCAGCTAAGAGGCTTGTGGAGGCTGTGTA 775  
QY 696 CTGGCCACACTCAACATCTCTCTGAGCAGGAGCCATTAATCCGCTTACATGGGCGC 755  
DB 776 CTGACGACACTCAATTAATCTCTCTGAGCAGGAGCCGCTTATCTTCATATAGGGAGC 835  
QY 756 CCACCATCTCATGAGAGAGATGACCTTGTGGCTCAACGAACTCACTTGTAGG 815  
DB 836 CCACCATCTCATGAGAGAGATGACCTTGTGGCTCAACGAACTCACTTGTAGG 895  
QY 816 CTGTGCTATGATGAGAGTGTGGCAAGCTTCAAGGGGGACGACGAAAGTCTACTT 875  
DB 896 CTGTGCTATGATGAGAGTGTGGCAAGCTTCAAGGGGGACGACGAAAGTCTACTT 955  
QY 876 CTGTGAGGAGGAGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGG 935  
DB 956 CTGTGAGGAGGAGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGG 1015  
QY 936 GGGCCGTGTGCAAGAGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGG 995  
DB 1016 GCGAGAGTGTGTAAGGTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 1075  
QY 996 GTTCTGAGAGGCGGCTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1055  
DB 1076 GTTCTGAGAGGCGGCTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1135  
QY 1056 GAGGCGATGACAGCCCTGAGAGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1115  
DB 1136 GAGGCGATGACAGCCCTGAGAGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1195



QY 1116 TCAGCAGAGTGGGGTGGACATGCTGTCGCCATCTGTGATACCAGTTGGAGAGAT 1175  
 1196 TCAGCGCATGGGGGATATGAGACCTGCTGAGTTGTGATACAGTTGGAGAGAT 1255  
 QY 1176 CCAGCGGGTGGTGGAGGGCCCTATAGAGATACATGAGGAG--CCGGAAGTGGGAC 1233  
 1256 CCAGCAGAGTGGTGGAGGGTCCCTACAAAGAGTACAGAGGCGCCAGAAAGTGGGCG 1315  
 QY 1234 CCGTACACTGACCCCTTACCCAGCCCTGGCGCTGGCTGATTAACAATGGCATCGG 1293  
 1316 CCGTAACTGACCCCGGTACCCAGCCCTGGCGCTGGCTGATTAACAATGGCATCGG 1375  
 QY 1294 CCGCAGCGGTACACACAGTCCCTGG--AGCTACCCGACAAATCTCAACTCTCAAGAA 1352  
 1376 GACAAATGGGTACACAGTCCCTGGAGAGTGGCGGACAAACCTCAACTCAAGAA 1435  
 QY 1353 GACACCGGTGATGGAGAGAGAGTGGGGCTCGGTGGAGAGCGCCCTGCTGTGAAGA 1412  
 1436 GACACCGGTGATGGAGAGAGAGTGGGGCTCGGTGGAGAGCGCCCTGCTGTGAAGA 1495  
 QY 1413 GGGCAGCAACTTCAACCCAGCTGGTGGCGGACCGGTTACAGACTTGTATGAGGACCTTA 1472  
 1496 GAACATTAATTAACACAGAGTGGTGGCGGACAGAGGTCGCCAGGCTTGTATGAGGACCTTA 1555  
 QY 1473 TACAGTGTCTTCAATTTGGACAGAGAGAGCGGCTGGCTCAAGGCTGTGAGCCTGGGGCC 1532  
 1556 TACAGTGTCTTCAATTTGGACAGAGAGAGCGGCTGGCTCAAGGCTGTGAGCCTGGGGCC 1615  
 QY 1533 CTGGGTTACCTTATGAGAGAGTGAAGTCTTTGACAGAGAGCCATGGAAGCTGTGT 1592  
 1616 CTGGGTTACCTTATGAGAGAGTGAAGTCTTTGACAGAGAGCCATGGAAGTGTGTGT 1675  
 QY 1593 GCTATCTCAGAGAGAGAGTGTCTTTGGCGGCTCGGCTCAAGTGTGTGAGAGTGTGC 1652  
 1676 GCTATCTCAGAGAGAGAGTGTCTTTGGCGGCTCGGCTCAAGTGTGTGAGAGTGTGC 1735  
 QY 1653 CGTGGCGAGTGTAAAGTATGCTCTGTGACAGAGTGTCTGTGCGCGGAGACCTTA 1712  
 1736 TCGGGCGAGTGTAAAGTATGCTCTGTGACAGAGTGTCTGTGCGCGGAGACCTTA 1795  
 QY 1713 TTGGCGCTGAGAGTGAACAGAGAGCGGCTGTGTGG--CCGTGGGTGGCATTTGGATC 1769  
 1796 CTGTGGCTGAGAGTGAACAGAGAGCGGCTGTGTGGCGGACAGAGTGTGTGGGTG 1855  
 QY 1770 TTACTGATGACAGATGTGATGACCTGCGACACTTTCAGCAATTTGCAACCTCGGTGAG 1829  
 1856 CTTTGTGATGACAGATGTGATGACCTTTCAGCAATTTGCAACAGATTTGAGCAT 1915  
 QY 1830 TAAGATACAGTCAAGCGCACTTCCCAAAACATCAGCGTGTGGCGGACAGAGCTGTGT 1889  
 1916 TAA--AAAAGTCAAGTCTATTCCTCAAAACATCAGCGTGTGTGGCGGACAGAGCTGTGT 1972  
 QY 1890 GCTGCCCTGACAGTCTCTCTCAACTTGGCGCTGGCGGACTCCAGACCCCG 1939  
 1973 CTTACCTGCGACCTCTCTCTCAACTTGGCGGACTCCAGACCTTCG 2022  
 Db

RESULT 2  
 LOCUS BX370082 977 bp mRNA linear EST 08-MAY-2003  
 DEFINITION BX370082 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone cS0D1074YD21 5-PRIME, mRNA sequence.  
 ACCESSION BX370082  
 VERSION BX370082.1 GI:30441574  
 KEYWORDS EST.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 977)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 1123.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=cS0BAG0362C10.CS03435.1cluster=1123.r.  
 Contact: Feng Liang Email: fliang@life.com  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID: C50BAG0362C10.CS03435.1.  
 FEATURES  
 source  
 1. 977  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1074YD21"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(4T)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 204 a 289 c 277 g 207 t  
 ORIGIN  
 Query Match 36.48; Score 832.4; DB 13; Length 977;  
 Best Local Similarity 93.96; Pred. No. 1.6e-181;  
 Matches 900; Conservative 0; Mismatches 51; Indels 7; Gaps 3;  
 QY 496 TTCACTTCAATCCGCTCTCTGAGAGCCCTTACAAATGCTCCAGCAGTACGCTGGGAGC 555  
 1 TTCACTTCAATCCGCTCTCTGAGAGCCCTTACAAATGCTCCAGCAGTACGCTGGGAGC 60  
 QY 556 TAGCCCTTCAGCCCAAGTGCACCTAGTCAACAGTCTCACTTGGAGCATGGA 615  
 61 TAGCCCTTCAGCCCAAGTGCACCTAGTCAACAGTCTCACTTGGAGCATGGA 120  
 QY 616 GAGTTTAAGATGGGAAGGGCAAGTGTCCCTATGAGAGCCAGTAAAGGCGCATGTGGCTT 675  
 121 GAGTTTAAGATGGGAAGGGCAAGTGTCCCTATGAGAGCCAGTAAAGGCGCATGTGGCTT 180  
 QY 676 CTTGTGATGATGAGTGTGATGAGTGTGAGAGAGTCAACAACTTCTGGGACAGGAACCAT 735  
 181 CTTGTGATGATGAGTGTGATGAGTGTGAGAGAGTCAACAACTTCTGGGACAGGAACCAT 240  
 QY 736 ATCTGCGTAACATGGGGGCCCCACCACTCATGAAAGAGAGTACCTGGGCTTTGGCTC 795  
 241 ATCTGCGTAACATGGGGGCCCCACCACTCATGAAAGAGAGTACCTGGGCTTTGGCTC 300  
 QY 796 AAGCAACTCACTTTGAGGCTGTGCTATGATGATGAGTGTGGGACAGTTCACGGGG 855  
 301 AAGCAACTCACTTTGAGGCTGTGCTATGATGATGAGTGTGGGACAGTTCACGGGG 360  
 QY 856 GAGCAGCACAAGGCTCTCTCTTCTTCAAGGAGGAGGAGGAGTGGAGTGGCATGCTATAGCC 915  
 361 GAGCAGCACAAGGCTCTCTCTTCTTCAAGGAGGAGGAGTGGAGTGGCATGCTATAGCC 420  
 QY 916 GAGCAGGATGTGCTGTGAGCCCGTGTGCAAGAGGAGATATGAGGGGCGCAGGAGCC 975  
 421 GAGCAGGATGTGCTGTGAGCCCGTGTGCAAGAGGAGATATGAGGGGCGCAGGAGCC 480  
 QY 976 CTGCAAGAGGAAGTGGACAGTTCCTTAAGCGCGGCTGGCATGCTTGCCTCCCAACTGG 1035  
 481 CTGCAAGAGGAAGTGGACAGTTCCTTAAGCGCGGCTGGCATGCTTGCCTCCCAACTGG 540  
 QY 1036 CAGCTTACTTCAACAGCTGACAGGAGATGACACCTTGGAGAGAGACCTCTGGCACAAC 1095  
 541 CAGCTTACTTCAACAGCTGACAGGAGATGACACCTTGGAGAGAGACCTCTGGCACAAC 600  
 QY 1096 ACCACCTTCTTTGGGGTTTCAAGCAGAGTGGGGTGAATGATGATGCTGGGCGCATCTGT 1155

Db 601 ACCACCTCTTGGGCTTTTAAACACAGTGGGGTGACATGTACCTGTGGCCATCTGT 660  
QY 1156 GAGTACGATTGGAAGATCCAGCGGGTGTGTGAGGGCCCTATAGAGATACCATGAG 1215  
Db 661 GAGTACGATTGGAAGATCCAGCGGGTGTGTGAGGGCCCTATAGAGATACCATGAG 720  
QY 1216 GAGCGCCGAGAGTGGGACCGCTACACTGACCTGTACCGAGCCCTGCGCTGCTGTC 1275  
Db 721 GAGCGCCGAGAGTGGGACCGCTACACTGACCTGTGCGCCCTGCGCTGCTGTC 780  
QY 1276 ATTACCACTGGCATCGGCGCCAGGCTACACCACTGCTGCGACTGACCCGACCAATC 1335  
Db 781 AATTAACAATGCGATCGGCGCCAGCTGTACACCACTGCTGCGACTGACCCGACCAATC 840  
QY 1336 CTCACCTGCTGTCAGAGACACCGCGCTGATGAGAGAGTGGGCTGCTGCTGAGCCGC 1395  
Db 841 CTAAATTTGTTAAGAA---CAACCCGCTGATGAGAGAGAGTGGGCTGCTGCTGAGCCG 897  
QY 1396 CCCCTGCTGCTGAGAGAGGACCAACTTACACCACTGCTGCTGCGCCGAGCGGTTACAG 1453  
Db 898 GCCCTGCTGCGGA--AAAGGCGACCAACTT--ACCCCTGCGGGGCGCGCCGCTCACGG 951

## RESULT 3

LOCUS BX369159 932 bp mRNA linear EST 08-MAY-2003

DEFINITION BX369159 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS001027YN09 5-PRIME, mRNA sequence.

ACCESSION BX369159

VERSION BX369159.1 GI:30455790

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 932)

L1, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 1123.r for

more information about this cluster, see

http://www.genoscope.cns.fr/

cg1-bln/cluster.cgi?seq=CS0BAF0242B08.AF02267.2&cluster=1123.r.

Contact: Fong Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID: CS0BAF0242B08.AF02267\_2.

Location/Qualifiers

1..932

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1027YN09"

/label\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 195 a 269 c 275 g 190 t

ORIGIN

Query Match 36.4% Score 831.8 DB 13: Length 932;

Best Local Similarity 95.3% Pred. No. 2,2e-181;

Matches 888; Conservative 0; Mismatches 40; Indels 4; Gaps 3;

Db 1 CTTCCAGCCCAAGTGCACCTACCTAACAATGAGCTCACTTATTGGAGCATGAGAGTT 60  
QY 621 TGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGTAAAGGCGCATGTGCGCTTGT 680  
Db 61 TGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGTAAAGGCGCATGTGCGCTTGT 120  
QY 681 GATGATGAGTGTACTGTGCGGACACTCAACACTTCTGCGGACAGGAACCATTAATCT 740  
Db 121 GATGATGAGTGTACTGTGCGGACACTCAACACTTCTGCGGACAGGAACCATTAATCT 180  
QY 741 GCTGATGAGTGTGCGGACACTCAACACTTCTGCGGACAGGAACCATTAATCT 800  
Db 181 GCTGATGAGTGTGCGGACACTCAACACTTCTGCGGACAGGAACCATTAATCT 240  
QY 801 ACCCTCACTTGTAGGCTGTGCTATGATGATGAGTGTGCGGACACTCAACAGGAGAGA 860  
Db 241 ACCCTCACTTGTAGGCTGTGCTATGATGATGAGTGTGCGGACACTCAACAGGAGAGA 300  
QY 861 CGACAAAGTCTACTTCTTCTGAGGAGGCGGCACTGAGTCCGACTGTATGCCAGCA 920  
Db 301 CGACAAAGTCTACTTCTTCTGAGGAGGCGGCACTGAGTCCGACTGTATGCCAGCA 360  
QY 921 GGTGATGAGTGTGCGGACACTCAACACTTCTGCGGACAGGAACCATTAATCT 980  
Db 361 GGTGATGAGTGTGCGGACACTCAACACTTCTGCGGACAGGAACCATTAATCT 420  
QY 981 GAGGAAGTGGACACAGTCTCTGAGGCGGCGTGGCATCTCTGCCGCAACTGGCAGCT 1040  
Db 421 GAGGAAGTGGACACAGTCTCTGAGGCGGCGTGGCATCTCTGCCGCAACTGGCAGCT 480  
QY 1041 CTACTTCAACACAGCTGACAGGCGATGACACCTGCGAGGACACTCTGCGACACACAC 1100  
Db 481 CTACTTCAACACAGCTGACAGGCGATGACACCTGCGAGGACACTCTGCGACACACAC 540  
QY 1101 CTTCTTGGGGGTTTTCAGACACAGTGGGGGTGACATGTACTGTGCGGCACTGTGAGTA 1160  
Db 541 CTTCTTGGGGGTTTTCAGACACAGTGGGGGTGACATGTACTGTGCGGCACTGTGAGTA 600  
QY 1161 CCAGTTGGAAGAGATCCAGCGGGGTGTTGAGGGCCCTATTAAGAGTACATAGAGAAC 1220  
Db 601 CCAGTTGGAAGAGATCCAGCGGGGTGTTGAGGGCCCTATTAAGAGTACATAGAGAAC 660  
QY 1221 CCAGAAAGTGGGACCGCTACACTGACCTGTACCCAGCCCTGCGGCTGTGCTGATTA 1280  
Db 661 CCAGAAAGTGGGACCGCTACACTGACCTGTACCCAGCCCTGCGGCTGTGCTGATTA 720  
QY 1281 CAAGTGGCATGCGGCGCCACAGGCTACACCACTGCTGAGTACCCGACCAATCTCA 1340  
Db 721 CAAGTGGCATGCGGCGCCACAGGCTACACCACTGCTGAGTACCCGAC -ACATGCTCA 779  
QY 1341 CTTGCTCAAGAAAGACACCGCTGATGAGAGAGAGTGGGGCCTC--GGTGAAGCGGCC 1398  
Db 780 CTTGCTCAAGAAAGACACCGCTGATGAGAGAGAGTGGGGCCTC--GGTGAAGCGGCC 839  
QY 1399 CTGCTGCTGGAAGAGGCGACCACTTACACCTGCTGAGGCGGCTGATGAGAG -ACT 1457  
Db 840 CTGCTGCTGGAAGAGGCGACCACTTACACCTGCTGAGGCGGCTGATGAGAGACT 899  
QY 1458 TGATGAGACCACTTATACAGTCTGCTGATG 1489  
Db 900 TGATGAGACCACTTATACAGTCTGCTGATG 931

## RESULT 4

LOCUS B1518922 864 bp mRNA linear EST 29-AUG-2001

DEFINITION B1518922 603061844F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5210938 5', mRNA sequence.

ACCESSION B1518922

VERSION B1518922.1 GI:15343714

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens



OY	823	TATGTACCTGAGAGTGTGGGAGCTTCAACGGGGGAGCAGCAAGGCTTACTTCTTC	882
Db	1	TTTGTCCCTGAGAGGTGTGGAAAGCTTCAACGGGAGACGATGACAGATCTACTTCTTTC	60
OY	883	AGGAGCGGGGAGTGGAGTCCGACTGCTATGACCAGAGAGTGTGGCTGTGTGGCCGT	942
Db	61	AGTAGCGGGGACGTGGAGTATGACTGCTATTTCGAGAGGTGTGGCTGTGTGGCCAGA	120
OY	943	GTCGTGCAAGGCGCATATGCGGGGGGCGACAGAACCCGTCAGAGAGTAAGTGGACACGTTCTGTG	1002
Db	121	GTCGTGCAAGGCGCATATGCGGGGGGAGACAGAGCGCTCGAGAAAGTAAGACGAGCGTTCTGTG	175
OY	1003	AAGCGCGGGGATGGATGCTGTGCCCCGCAACAGTGGAGCTACTATTCACACAGCTGCAAGCG	1067
Db	176	AAGCGTGGGTTGTGTGTGCTCAAGCCCTGTACGTGGAAGGTCTACTTCACACAGCTGCAAGCG	235
OY	1063	ATGCACACCCCTGCAGAGACACTCTCTGCGACAAACACACCTTCTCTTGGGTTTTCAGCA	1122
Db	236	GTGCACACCCCTGCAGAGAGCGCTCTTGGGACAAACACACCTTCTCTTGGGTTTTCAGCG	295
OY	1123	CAGTGGGGTGACATGTACCTGTGCGGCCATCTGTAGTACCATGTTGGAGAGATCCAGCG	1187
Db	296	CGATGGGGCGATGTGAGACTGTCTGTGCAATTTGTGATGACAGTTGGAAACAATCCAGCAA	355
OY	1183	GTTGTTTGAAGGGCCCCCTATTAAGAGTACCATTAGAGAAAGCCAGAAAGTGGAGCCGCTACACT	1247
Db	356	GTTGTTTGAAGGGTCCCTACAAAGAGTACAGTGAAGCAAGCCAGAAAGTGGGCCCGGTATCT	415
OY	1243	GACCTGTATCCAGACGCTCTGGGCTGTGCTGTGATTAACAATGGGCATCGGCGCACAGGC	1307
Db	416	GACCTGTATCCAGACGCTCTGGGCTGTGCTGTGATTAACAATGGGCATCGGCGCACAGTGC	475
OY	1303	TACACAGAGCTCCCTGAGACTACCCGACAAATCCGCAACTTGTCAGAAAGACACCCGGTG	1362
Db	476	TACACAGAGTCCCTGAGAACTGCGGACACACACCTCAACTTCATCAGAAAGACACCCGTCG	535
OY	1363	ATGAGAGAGCAGGTGGGGCTCTGGGTGAGAGCCGCCCTGCTGTGTAAGAGGACCAACAC	1422
Db	536	ATGAGAGAGCAGGTGCAAGCCTCTGGGTGGGGCCGCCCTTACTTGTCAGAAAGACACCTAAC	595
OY	1423	TTTCAACCAACCTGGTGGGCGCGACCGGGTTTACAGAGACTTATATGAGAGCACTATACAGTGTG	1487
Db	596	TTTCAACCAACCTGGTGGGCGCGACAGGGTCCCAAGGGCTGTGATGGTGCACCTATACAGTGTG	655
OY	1483	TTTCATTGGCAGACAGAGAGAGCGCTGGCTGCTCAAGAGCTGTGAGACCTGTGGGCCCTGGAGTTAC	1542
Db	656	TTTCATTGGTACAGAGAGATGGCTGGCTGCTGAAGGCTGTGAGACCTGTGGGCCCTGGAGTTAC	715
OY	1543	CTGATTGAGAGAGCTGCAGCTGTGTTGACCAAGAGCCCATGAGAAAGCTGTGCTATATCTCAG	1602
Db	716	ATGCTGCAAGAGAACTGCAGAGTGTGTTGACACAGAGACAGTGAAGAGTCTGTGCTGTCTCAG	775
OY	1603	AGCAAGAAAGCGTCTTTTGGCGGGTCCCGGCTCTACAGCGGTGTGAGAGCTGCCGTGGCCGAC	1662
Db	776	AGCAAGAAAGGCGTCTTTTGGCGGGTCCCGGCTCTACAGCGGTGTGAGAGCTGTGCTGGCCGAC	835
OY	1663	TGCATTAAGATATCG-CTCCTGTGTGACAGACTGTCTCTGCGCCGCGAAGCCCTATATGCG	1716
Db	836	TGCACAAAGTACGCTTTTCTGTGTGTAAGACTGTCTCTGTGGNAGAGGGAGACCCTTAC	890
RESULT 6			
BE294211			
LOCUS	60117274.F1	705 bp	mrna
DEFINITION	NIH_MGC_17	Homo sapiens	cdna clone IMAGE:3528227 5', linear EST 20-JUL-2000
ACCESSION	BE294211		
VERSION	BE294211.1		
KEYWORDS	EST		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE AUTHORS TITLE JOURNAL COMMENT	FEATURES SOURCE
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 705) NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgsb@rsfemail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">image.lnl.gov</a> Plate: L1CML197 row: 9 column: 12 High quality sequence stop: 624. Location/Qualifiers 1..705 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone IMAGE:3528227" /tissue_type="rhabdomyosarcoma" /lab_host="DH10B (phage-resistant)" /clone_1lb="NIH-MGC_117" /note="Organ: muscle; Vector: pOTB7, site_1: EcoRI; site_2: XhoI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."	
BASE COUNT	143 a 209 c 215 g 138 t
ORIGIN	
Query Match	27.7% Score 632.6; DB 10; Length 705;
Best Local Similarity	98.2% Pred. No. 2,1e-135;
Matches	661: Conservative 0; Mismatches 9; Indels 3; Gaps 2;
QY	749 TGGGGCCCCACCACTCCATGAGAGACAGAGTACCTGGCCCTTTTGGCTCAAGCAACCTCACT 808
Db	1 TGGGGCCCCACCACTCCATGAGAGAGAGTACCTGGCCCTTTTGGCTCAAGCAACCTCACT 60
QY	809 TTGAGGCTCGCTATGTAACCTGAGAGTGTCGGACCTTCAGCGGGGAGCAGACAAAG 868
Db	61 TTGAGGCTCGCTATGTAACCTGAGAGTGTCGGACCTTCAGCGGGGAGCAGACAAAG 120
QY	869 TCTACTCTCTTTCAGGGAGCGGCGACATGTGACGACTGCTATGCCAGACAGGTGGTGG 928
Db	121 TCTACTCTCTTTCAGGGAGCGGCGACATGTGACGACTGCTATGCCAGACAGGTGGTGG 180
QY	929 CTCGTGTGGCCCGCTGTCTGCAGAGGCGCATATGGGGGGCGCACAGCACTTCAGAGGAAT 988
Db	181 CTCGTGTGGCCCGCTGTCTGCAGAGGCGCATATGGGGGGCGCACAGCACTTCAGAGGAAT 240
QY	989 GGACCAAGTTCCTAAGAGCGCGGTGACATGCTGTGCCCGAAGCTGGCAGCTCACTTCA 1048
Db	241 GGACCAAGTTCCTAAGAGCGCGGTGACATGCTGTGCCCGAAGCTGGCAGCTCACTTCA 300
QY	1049 ACCAGCTGCAGAGGATGCACACCTGACAGACCTCTGGCAGACACCAACCACTTCTTGG 1108
Db	301 ACCAGCTGCAGAGGATGCACACCTGACAGACCTCTGGCAGACACCAACCACTTCTTGG 360
QY	1109 GGGTTTTTCAAGCAAGTGGGGGTACATGTACCTGTGGGCGCACTGTGAGATACCAATGG 1168
Db	361 GGGTTTTTCAAGCAAGTGGGGGTACATGTACCTGTGGGCGCACTGTGAGATACCAATGG 420
QY	1169 AAGAGATCCAGCGGGGTGTTGAGGGGCCCTATAAGAGTACATGAGAGAGCCAGAAAT 1228
Db	421 AAGAGATCCAGCGGGGTGTTGAGGGGCCCTATAAGAGTACATGAGAGAGCCAGAAAT 480
QY	1229 GGGAGCGGTACACTGACCCCTGTACCCAGCCCTGGCTGGCTGCTGATTAACCACTGGC 1288

DB	481	GGGACCGCTACAATGAGCCCTGTACCACGACCCCTCGGCCCTGCCTGCAATTAAACAACAGGC	540
OY	1289	ATCGGCGCCACAGGCTACACCAAGTCCCTTG-CAGCTACCCGACACAAATCCTTAACTTGTC	1347
Db	541	ATTCGCGCGCACAGGGCTACACCAAGCTCCCTGTAAGCTACCCGACAAATCCTTAACTTGCTC	600
OY	1348	AAGAAGCACCGCGGTGATGGAGGACAGGTGGGGGCTCGGTGAGACCGCCCCCTGCTGTG	1407
Db	601	AAGAAGCACCGCGGTGATGGAGGAGAGAGGTGGCTCGGTGGAGAGGCCCC--TGTCTGTG	658
OY	1408	AAGAAGGCGCACCA 1420 	
Db	659	AAGAATGGCCCAA 671	
RESULT 7	BUS18484		
LOCUS	BUS18484	937 bp	mRNA linear EST 12-SEP-2002
DEFINITION	AGNCNCOVRT_10163679 NIH_MGC_134 Mus musculus cDNA clone		
IMAGE:	6516221 5' , mRNA sequence.		
ACCESSION	BUS18484		
VERSION	BUS18484.1		
KEYWORDS	EST .		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
TAXONOMY	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 937) NIH-MGC http://mgi.ncl.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT	Contact: Robert Strausberg, Ph.D.		

```

FEATURES
  source
    1. .937
    location/Qualifiers
    High quality sequence stop: 671.
    High quality sequence start: 19
    Plate: LLNMI4094
    row: 9
    column: 06
    http://image.llnl.gov
    found through the I.M.A.G.E. Consortium/LLNL at:
    Clone distribution: MGC clone distribution information can be
    DNA sequencing by: Agencourt Bioscience Corporation
    cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    cDNA library Preparation: Invitrogen Corp

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_1="IMAGE:6516221"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC.134"
/notes="Vector: pCMV-Sport6.1.ccd; Site_1: EcoRV; Site_2:
NotI; Cloned unidirectionally. Primer: Oligo dh. Average
Insert size 1.7 kb. Constructed by Resgen, Invitrogen
Corp. Note: this is a NIH_MGC library."

```

Query Match	Match Similarity	Score	DB	Length
Best Local	85.48	Pred. No. 9	133	
Matches	703	Conservative	0	Matches 119; Indels 1; Gaps 1
QY	GGGAGTGGAGTCCGACTGCTATGTCGAGAGAGGTGGTGGCTCGTGTGGCCGCTGTCTCA	949		
DB	GGGAGTGGAGTATGCTCTATTCGAGAGAGGTGGTGGCTCGTGTGGCGAGTCTGTCA	86		
QY	AGGGGATATGGGGGGGCGAGCGAGCCGTCAGAGGAAGTGGACCACTTTCGGAAGGGGC	1005		
DB	AGGGGATATGGGGGGGCGAGCGAGCCGTCAGAGGAAGTGGACCACTTTCGGAAGGGGC	1005		
QY	AGGGGATATGGGGGGGCGAGCGAGCCGTCAGAGGAAGTGGACCACTTTCGGAAGGGGC	1005		
DB	AGGGGATATGGGGGGGCGAGCGAGCCGTCAGAGGAAGTGGACCACTTTCGGAAGGGGC	1005		
QY	GGCTGGCATGCTCTGCCCCGAAGTGGAGCTTACTTCAACACAGTGCAGGGGATGACA	1065		
DB	GGCTGGCATGCTCTGCCCCGAAGTGGAGCTTACTTCAACACAGTGCAGGGGATGACA	1065		

Db	147	GGTTGGTGTGCTCAGGCCCTTACGTGGAAAGTCTACTTTCACACCACTGTAAAGCGGTGCACA	206
QY	1070	CCCTGCAGGACACCTCCTGTGCACACACACACCTTCTTTGGGGTTTTTCAMGACATGTGG	1129
Db	207	CCCTTCGGGGGGCGCCCTCTTGGGCACAAACACACCTTCTTCGGGGTTTTTCAMGCGGATGG	266
QY	1130	GTGACATGTACCTGTCCGGCCATGCTGTGTACAGTTGGGAAGATATCAGCGGGTGTGTTG	1189
Db	267	GGGATATGTGACTGTCTGTCACTTTTGTGTGTACCATTTGTGAAACATATCCAGTAAGTGTGG	326
QY	1190	AGGGCCCTATTAAGAGATACCATGTGAGAGGCCAGAAAGTGGAGCCGTACACTGACCTGT	1249
Db	327	AGGGTCCCTACAAAGGAGATACAGTGTGAGCAAGCCCAAGTGGGCCCGCTATCTACGCCGG	386
QY	1250	TACCCAGGCCCTCGGCTGGCTCGTGCATTAACAATCTGGCATTCGGCGCAGCGGTATACACA	1309
Db	387	TACCCAGCCCTCGGCTGGCTGTGTGTATCAACAATCTGGCACCGAGACATGTATCACCA	446
QY	1310	GCTCCCTGGAGGTACCCGACCAACATCTCAACTTCGTGTAAGAACACCCCGCTGATGTGAGG	1369
Db	447	GTTCCCTGGAACTGGCCGGACAAACCCCTCAACTTCATCAAGAACACCCCGCTGATGTGAGG	506
QY	1370	AGCAGGTGGGCTCTGTGTGTGAGCCGCCCTCTGCTGTGTGAAAGAGGCACCACTTCACCC	1429
Db	507	ACCAGGTAAAGCCTCGTGTGTGGGCCGCCCTCTACTGTGTGAAMAACAACTCAACTTCACAC	566
QY	1430	ACCTGTGTGGCCAGCCGGGTTACAGAGATCTGTATGAGGCCACCTATACAGTGTCTTCATGT	1489
Db	567	ACGTGTGTGGCCAGCAAGGGTCCCAAGGGCTGTATGTGTGTGCACCTATACAGTGTGTTCATGT	626
QY	1490	GCACAGAGAACCGTGTGCTCTCAAGGCTGTGAGCGCTGGGGCCCTCTGGGTTCACTGATTG	1549
Db	627	GTACAGGAGATAGGCTGTGCTGTGTGAAGCCTGTGAGGCTGTGCGCTGTGATCCAGATGTGG	686
QY	1550	AGGACTGTGAGCTGTTTGTACCGAGAGGCCATAGAGAAAGCTGTGTGTATCTCAGAGACAGA	1609
Db	687	AGGAATCTGCAAGTGTTTGTACCGAGAGGCCATGTGAAGAAAGTCTGTGTGTCTGTAGACACGA	746
QY	1610	AGCTGCTCTTTGCCGGCTCCGCGTTCAGTGTGTGACAGTCCCGGTGGCCGACTGTCAATA	1669
Db	747	AGGTGCTCTTTTCTGTGGCTCCGCGCTCTCAGCTGTGTCACTGTCTCTCTGTGGCAGCTGCACA	806
QY	1670	AGCTA-TCCGCTCTGTGACAGATGTGTCCGCGCGGAGCCCT 1711	
Db	807	ACTACCCGTTTCTGTGTAACTGTGTCTCTGGCCAGGAACCTT 849	

RESULT	B
Locus	BM926974
DEFINITION	AGENCOURT_6682051 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5767867 5', mRNA sequence.
ACCESSION	BM926974
VERSION	GI:19377353
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1394) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nhl.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Clone distribution: Agencourt Bioscience Corporation found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM12826 row: m column: 20

High quality sequence stop: 574.

FEATURES  
Location/Qualifiers  
1..1394

source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5767867"  
/lab\_host="DH10B"  
/clone\_11b="NIH\_MGC\_121"  
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."  
BASE COUNT 373 a 438 c 338 g 244 t 1 others  
ORIGIN

Query Match 26.4%; Score 603.8; DB 12; Length 1394;  
Best Local Similarity 99.7%; Pred. No. 1.2e-128;  
Matches 605; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1154 GTGACTACAGTGGAGAGATCCAGCGGCTGTTGAGGGCCCTATAGAGATACATG 1213  
1 GTGACTACAGTGGAGAGATCCAGCGGCTGTTGAGGGCCCTATAGAGATACATG 60  
1214 AGGAGCCAGAGTGGAGAGATCCAGCGGCTGTTGAGGGCCCTATAGAGATACATG 1273  
61 AGGAGCCAGAGTGGAGAGATCCAGCGGCTGTTGAGGGCCCTATAGAGATACATG 120  
1274 GCATTACAGTGGAGAGATCCAGCGGCTGTTGAGGGCCCTATAGAGATACATG 1333  
121 GCATTACAGTGGAGAGATCCAGCGGCTGTTGAGGGCCCTATAGAGATACATG 180  
1334 TCCTGACCTGCTCAAGAACACCGCGTGATGAGAGAGAGAGAGAGAGAGAGAG 1393  
181 TCCTGACCTGCTCAAGAACACCGCGTGATGAGAGAGAGAGAGAGAGAGAGAG 240  
1394 GCCCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1453  
241 GCCCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
1454 GACTTATGAG 1513  
301 GACTTATGAG 360  
1514 AGGCTGTAG 1573  
361 AGGCTGTAG 420  
1574 AGCCATGAG 1633  
421 AGCCATGAG 480  
1634 CTCACCTGCTGAG 1693  
481 CTCACCTGCTGAG 540  
1694 TCCTGAG 1753  
541 TCCTGAG 600  
1754 GTGGCCA 1760  
601 GTGGCCA 607

RESULT 9  
BM943605

LOCUS BM943605 768 bp mRNA linear EST 14-MAR-2002  
DEFINITION UI-M-EH0P-bvn-b-09-0-UI.r1 NIH\_BMAP\_EH0P Mus musculus cDNA clone  
IMAGE:5694248 5', mRNA sequence.  
ACCESSION BM943605  
VERSION BM943605.1 GI:19427190  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE NIH-MGC <http://mgi.ncl.nih.gov/>.  
1 (bases 1 to 768)  
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)  
UNPUBLISHED  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP).  
Seq primer: pYX-5.  
FEATURES  
source  
1..768  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5694248"  
/tissue\_type="Whole brain"  
/dev\_stage="embryo 18.5 dpc"  
/lab\_host="DH10B (TI phase resistant)"  
/clone\_11b="NIH\_BMAP\_EH0P"  
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I; Site: 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GAGCAGAGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 176 a 196 c 218 g 177 t 1 others  
ORIGIN

Query Match 26.2%; Score 597.6; DB 12; Length 768;  
Best Local Similarity 86.9%; Pred. No. 2.6e-127;  
Matches 668; Conservative 0; Mismatches 100; Indels 1; Gaps 1;

254 GTAAGACAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313  
1 GGAAGACAGTATCTTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
314 TCACAGACTCTCTGACAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 373  
61 TCACAGACTCTCTGACAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
374 GAGAGCCCTGTTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433  
121 GAGAGCCCTGTTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
434 AGGCCCCGTGAG 493  
||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1 ||||| 1

Db 181 AGGCTCAGCTGAGAGAAATTTGATGTCACCGAAGAGGAGGAGACACACCAAT 240  
 QY 494 GCTTCAACTTCAATCCGCTTCTGACAGCCCTCAATGCTCCACCTGTACGTGTGCA 553  
 Db 241 GCTTCAACTTCAATCCGCTTCTGACAGCCCTCAATGCTCCACCTGTATGTGGGCA 300  
 QY 554 CTTACGCTTCCAGCCCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 613  
 Db 301 CTTACGCTTCCAGCCCACTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 360  
 QY 614 GATAGTTGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 673  
 Db 361 CAGAAATTTGAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
 QY 674 TTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733  
 Db 421 TTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 QY 734 TTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793  
 Db 481 TTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 794 TTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853  
 Db 541 TTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 QY 854 GGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 913  
 Db 601 GAGACAGTGCAGAGTGTACTTCTTCTGAGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 QY 914 CCGACAGT 973  
 Db 661 CCGACAGT 720  
 QY 974 CCTGTGAGAGAGT 1022  
 Db 721 CGCTGCAGAGAA-TGGACGACCTTCTGTGAGGCTGCTGTGTGTGTGTGTGTGTGT 768  
 RESULT 10  
 BE270240 632 bp mRNA linear EST 13-JUL-2000  
 LOCUS BE270240 NIH\_MGC\_8 Homo sapiens cDNA clone IMAGE:3543559 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE270240  
 VERSION BE270240.1 GI:9143872  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 632)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at: image.lmnl.gov  
 Plate: LCM237 row: f column: 08  
 High quality sequence stop: 631.  
 Location/Qualifiers  
 1. 632  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3543559"  
 /tissue\_type="Burkitt lymphoma"  
 /lab\_host="DH10B (phage-resistant)"

/clone\_1b="NIH\_MGC\_8"  
 /note="Organ: Lymph; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAAGAG(G). Size-selected >500bp for Hong  
 Insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 131 a 185 c 192 g 124 t  
 ORIGIN  
 Query Match 26.1%; Score 597; DB 10; Length 632;  
 Best Local Similarity 99.5%; Pred. No. 3.4e-127;  
 Matches 630; Conservative 0; Mismatches 0; Indels 3; Gaps 3;  
 QY 750 GGGGCCCCACCACTCCATGAAAGACAGATACCTGCTTTTGGCTCAACGACCTCACT 809  
 Db 1 GGGGCCCCACCACTCCATGAAAGACAGATACCTGCTTTTGGCTCAACGACCTCACT 60  
 QY 810 TGTAGGCTTGCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869  
 Db 61 TGTAGGCTTGCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 870 CTACTTCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 929  
 Db 121 CTACTTCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
 QY 930 TCGTGTGGCCCGT 989  
 Db 181 TCGTGTGGCCCGT 240  
 QY 990 GACACAGTTCCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1049  
 Db 241 GACACAGTTCCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
 QY 1050 CCAGCTCAGGCGATGACACACCTGTCAGAGACCTCTGTGGCAACACACACCTCTT 1109  
 Db 301 CCAGCTCAGGCGATGACACACCTGTCAGAGACCTCTGTGGCAACACACACCTCTT 360  
 QY 1110 GGTTTTCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1169  
 Db 361 GGTTTTCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
 QY 1170 AGAGATCCAGCGGCTGT 1229  
 Db 421 AGAGATCCAGCGGCTGT 480  
 QY 1230 -GGACCGCTACACTGACCTGTATCCAGCCCTGCGCTGTGTGTGTGTGTGTGTGTGT 1288  
 Db 481 TGGACCGCTACACTGACCTGTATCCAGCCCTGCGCTGTGTGTGTGTGTGTGTGTGT 540  
 QY 1289 ATGGGCGCCAGCGGCTACACAGCTCCCTGAGACTACCCGACACATCTCAACTGCTCA 1348  
 Db 541 ATGGGCGCCAGCGG-TACACAGCTCCCTGAGACTACCCGACACATCTCAACTGCTCA 599  
 QY 1349 AGAAGCAGCCGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1381  
 Db 600 AGAAGCA-CCGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 631  
 RESULT 11  
 B0068313 1002 bp mRNA linear EST 02-APR-2002  
 LOCUS B0068313  
 DEFINITION AGENCOURT-6794173 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5770567  
 5', mRNA sequence.  
 ACCESSION B0068313  
 VERSION B0068313.1 GI:19897359  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.



REFERENCE 1 (bases 1 to 1002)  
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>,  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgap@biml.nih.gov](mailto:cgap@biml.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LHM12833 row: n column: 08  
 High quality sequence stop: 689.  
 Location/Qualifiers  
 1..1002  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5770567"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb. Insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."  
 BASE COUNT 176 a 303 c 355 g 167 t 1 others  
 ORIGIN  
 Query Match 25.4%; Score 579.6; DB 13; Length 1002;  
 Best Local Similarity 95.1%; Pred. No. 4,1e-123;  
 Matches 619; Conservative 0; Mismatches 30; Indels 2; Gaps 2;  
 84 CAGGCCGGGACCGGACCTCGCTGCGACCGCGCGCACAGACAGACAGCTGCGCTGMA 143  
 143 CCGAAGCGGAGCGGGCGCGCGCGGAGCTGAGAGCTGGCGGCAATAGACACGCTGCTGA 202  
 144 GCTCAGACCGGGGGGCTGCGCCATGCGCCACACTGGGCTGTCTGGCTGGCAGCANG 203  
 203 GCTCAGACCGGGGGGCTGCGCCATGCGCCACACTGGGCTGTCTGGCTGGCAGCANG 262  
 204 GCTGAGGGGCGCTGGGCAATGGGGCTGAGTGTGTGGAACCTTGGCGCGTAACAGT 263  
 263 GCTGAGGGGCGCTGGGCAATGGGGCTGAGTGTGTGGAACCTTGGCGCGTAACAGT 322  
 264 GCTCTTCTGGGAGCTGGCCACGAGTACGGCGGCTTCTCCAGACCGGCAATCAGAGT 323  
 323 GCTCTTCTGGGAGCTGGCCACGAGTACGGCGGCTTCTCCAGACCGGCAATCAGAGT 382  
 324 CCTGACACTGACGCTGACGAGACCGCACTGGGCTTGTACGTGGGCGCCGAGAGGCTT 383  
 383 CCTGACACTGACGCTGACGAGACCGCACTGGGCTTGTACGTGGGCGCCGAGAGGCTT 442  
 384 GTTTCCTTTCAGCATGAGAGCGCTGGAGCTGCAAGAGCATCTCTGGAGAGCCCGCT 443  
 443 GTTTCCTTTCAGCATGAGAGCGCTGGAGCTGCAAGAGCATCTCTGGAGAGCCCGCT 502  
 444 GGAGAGAAGACTGAGTATCCAGAAAGGAGAAACAACAGACGAGAGTGTCAAGT 503  
 503 GGAGAGAAGACTGAGTATCCAGAAAGGAGAAACAACAGACGAGAGTGTCAAGT 562  
 504 CATCGCTTCTGACGCGCTTACATGCTCCCACTGTACGTCTGTGGCACTACGCTT 563  
 563 CATCGCTTCTGACGCGCTTACATGCTCCCACTGTACGTCTGTGGCACTACGCTT 622  
 564 CCAAGCAAGTGCACACTACGTCAACATGCTACCTTACCTTGGAGCATGGAGATTTGA 623

DB 623 CCAGCCCAAGTGCACCTACGTCAACATGCTACCTTTCAGAGCATGAGAGTTCGA 682  
 624 AGATGGGAAGGCAAGTGTCCCTATGACCCACTAAGGGCCATGTGG-CCTTCTTGG 682  
 DB 683 AGATGGGAAGGCAAGTGTCCCTATGACCCACTAAGGGCCATGTGG-CCTTCTTGG 742  
 683 ATGTGAGCTGTACTGCGCCACACTCAACA-CCTTCTGCGGACGGAACCC 732  
 743 ATGTGAGCTGTACTGCGCCACACTCAACA-CCTTCTGCGGCGGGAACCC 793  
 RESULT 12  
 BX396978 853 bp mRNA linear EST 13-MAY-2003  
 LOCUS BX396978 Homo sapiens PLACENTA cot 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS001027YN09 5-PRIME, mRNA sequence.  
 ACCESSION BX396978  
 VERSION BX396978.1 GI:30609298  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 853)  
 L1.W.B., Gruber, C., Jessee, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr); Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 1123.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS001027CG050Plcluster=1123.r>. Contact : Feng Liang Email : [fliang@life.com](mailto:fliang@life.com) URL : <http://fulllength.invitrogen.com/invitrogen/Corporation/1600>  
 Faraday Avenue Genoscope sequence ID : CS001027CG050Pl.  
 Location/Qualifiers  
 1..853  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS001027YN09"  
 /tissue="Placenta"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 145 a 253 c 283 g 159 t 13 others  
 ORIGIN  
 Query Match 25.2%; Score 575.4; DB 13; Length 853;  
 Best Local Similarity 97.8%; Pred. No. 3.6e-122;  
 Matches 576; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
 114 CGGCGGCAACAGAGACAGCTGGGCTGAAGCTCAACACCGGGGCTGGCCATGAGCTG 173  
 265 CGGCGGCAACAGAGCTGGGCTGAAGCTCAACACCGGGGCTGGCCATGAGCTG 324  
 174 ACAGTGGGCTGTCTGGCTGGCAGCAAGAGCTGGGCTGGGCTGGGCTGGAGT 233  
 325 ACAGTGGGCTGTCTGGCTGGCAGCAAGAGCTGGGCTGGGCTGGGCTGGAGT 384  
 234 GTGGTGAACCTTGTGGCGCGGTAAGACAGTGTCTTCTGGGAGCTGGCCAGCGTATAG 293  
 385 GTGGTGAACCTTGTGGCGCGGTAAGACAGTGTCTTCTGGGAGCTGGCCAGCGTATAG 444  
 294 GCGGTTTCCCGAGACGGGCAATCCAGAGCTTCTGACACTGACGCTGAGGAGCCACATGG 353



[illegible]

RESULT 13	LOCUS	DEFINITION
BM962964	BM962964	745 bp mRNA linear EST 18-MAR-2002 UI-M-800-bw1-c-05-0-ri r1 NIH.BMAP_800 Mus musculus cDNA clone IMAGE:569876 5', mRNA sequence. mw02024

ACCESSION	BM902504	GI:19546384
VERSION	BM962964.1	
KEYWORDS	EST	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 745)	NIH-MGC	<a href="http://mgc.nhl.nih.gov/">http://mgc.nhl.nih.gov/</a>	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished
	Contact: Robert Strausberg, Ph.D.			

Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/INL at:  
<http://image.jnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
Seq primer: pTX-5.

FEATURES	Location/Qualifiers
source	1. .745

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5698876"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10b (r1 phage resistant)"
/clone_lib="NH_BMAP_E00"
/notes="Organ: brain; Vector: pUX-Asc; Site:1: EcoR I; Site:2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction,

```

ligated with *Eco*R I adaptor, digested with *Not* I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the *Not* I site and the polyA tail, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), 'Hemlin Chin, Ph.D., program coordinator.'

Query Match	24.9%	Score	569.2	DB 121	Length	745
Best Local Similarity	85.6%	Pred. No.	9.4e-121			
Matches 631; Conservative	0	Mismatches	106		Indels	0
					Gaps	0

OY	425	TCTCCTGGGAGGCCCCCTGTGGAGAAGAACTGAATGCACAGAAAAGGAACAACC	484
Db	7	TCTTTTGGAGGCTCCACTCGTAGAAGAAATTGTAATGACCCAGAAAGGAAGGACCAAC	66
OY	485	AGACCGAATGCTTCACATTCATCCGCTTCCTGCAACCCCTAACAAAGCTCCCACTGTACG	544
Db	67	AGACCGAATGCTTCACATTCATCCGCTTCCTGCAACCCCTAACAAATTCCTCCCATCTGTATG	126
OY	545	TCTGTGCACTAGACCTTCCAGGCCAAGTGACACTAGCTCAACATGCTCACCTTCATT	604
Db	127	TCTGTGCACTATGCTCCATCCAGGCCAAGTGACACTAGCTCAACATGCTCACCTTCATT	186

605 TGGAGCAATGGAGAGTTTGAAGAGTGGGAAGGCAAGTGTCCCTATGACCCAGCTAAGGGCC 664

Db 187 TGGACCGTGCAGCATTTGAGGATGGGMAATGCCCATATGACCCAGCTAAGGGTC 246

Db 247 ACACCGACTCTTGTGGACGGTGGCTGACTACGCCACACTCAATTAACCTTCTGGGGA 306

CGGACCCATTATCTGCGTAACATGGGGCCACCACCTCCATGAGACAGAGTACCTGG 784

Db 307 CAGAGCCGGTTATCCTTCGATACATGGGGACCCACCACTCCATCAGACAGAGTACCCTGG 366

Dy 844  
CTTTTGGCTCAACGACCCTCATTGTAGSCTCTGCGCTATGCTACCCGAGAGCTGGCGA

Db 426  
CTTTTGGCTGAATGTAACCCCACCTTGTAAGCTCTGCGCTTTGTGCCCGAGAGCTGGGAA

845 GCTTACGGGGGACGACGACAGGCTACTTCTTCTCAGGGAGCGGGCAGTGGATCG 904

Db 427 GCTTCAAGGAGACGATGACAAGATCTACTTCTTCACTGAGCGGCGACGTGGAGTATG 486

QY 905 ACTGCTATGCCAGCAGGTCGTGGCTCGTGTGGCCCGTGTGTGCAAGGGCGATATGGGG 964  
 Db 487 ACTGCTATTCGAGCAGGTGGTGGCTCGTGTGGCCGAGTGTGTAAGGGTGAATGGGG 545  
 |||||

965 GCGCAGCGACCCGTCAGAGGAAGTGACACAGTTCCTGAAGCGCGGCTGGCATGCTCG 1024

Db 547 GAGCACGGACGCTGCAGAGAAATGGACGACGTTCTGTAAGGCTCGTTGTGTCTCAG 606

Qy 1025 CCCCAGCTGGCAGCTCTACTTCAACCAAGTGCAGGCGATGCACACCTCGAGGACACT 1084

1085 CCTGGCACACACCACCTTCTTTGGGGTTTTCAGACACAGATGGGGTGACATGTACCTGT 1144

Db 667 CTTGGCACAACACCCTTCTCTGGGGTTTTCANCGCGATGGGGCATATGACCTGT 726

QY	1145	CGGCCATCTGTGAGTAC	1161
DB	727	CTGGCATCTGCTGCTAC	743

121 C10C001191000100 143

RESULT 14	
BU612234	
LOCUS	BU612234
	763 bp
MRNA	linear
	EST 30-FEB-2003

**DEFINITION** UI-M-EMO-cax-k-14-0-01.r1 NIH-BMAP-EMO Mus musculus cDNA clone UI-M-EMO-cax-k-14-0-01 5', mRNA sequence.

ACCESSION BU612234  
 VERSION BU612234.1 GI:23278437  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerogonathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bhs-remail.nih.gov  
 Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
 Seq primer: pYX-5.  
 Location/Qualifiers  
 1..763  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-EMO-cax-k-14-0-UI"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 15.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_id="NIH-BMAP\_EMO"  
 /note="Organ: brain; Vector: pYX-Asc; Site:1: EcoR I; Site:2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCCGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 171 a 207 c 222 g 163 t  
 ORIGIN

Query Match 24.3%; Score 554.6; DB 13; Length 763;  
 Best Local Similarity 86.2%; Pred. No. 2.2e-117;  
 Matches 625; Conservative 0; Mismatches 99; Indels 1; Gaps 1;

634 GGCAAGTCCCTATGACCCAGCTAAGGCCATGCTGCTCTTGTGATGCTAGCTG 693  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1 GGTAAATCCCATATGACCCAGCTAAGGCCATGCTGCTCTTGTGATGCTAGCTG 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 694 TACTGCGCACACTCAACCACTCTCTGGGACAGGAACCATTAATCTCGTAACTGGGG 753  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 61 TACTGCGCACACTCAACCACTCTCTGGGACAGGAACCATTAATCTCGTAACTGGGG 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 754 CCCGACACCTCATGAAGAGAGTACCTGGCTTTGGCTCAACGAACCTCACTTTGTA 813  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 121 ACCGACACCTCATGAAGAGAGTACCTGGCTTTGGCTCAACGAACCTCACTTTGTA 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 814 GGCCTCTGCTATGACCTGAGAGTGTGGGACCTTACAGGGGAGAGCAAGAGTCTAC 873  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 181 GGCCTCTGCTATGACCTGAGAGTGTGGGACCTTACAGGGGAGAGCAAGAGTCTAC 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 874 TTCTCTTCAGGAGCGGAGTGTGAGTGCAGTGTCTATGCGAGCAGTGTGCTGCTGT 933  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 241 TTCTCTTCAGGAGCGGAGTGTGAGTGCAGTGTCTATGCGAGCAGTGTGCTGCT 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 934 GTGGCCCGGTCTGCAAGGGCGCATATGAGGGGGCGACAGCCCTGACAGAAAGTGACC 993  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 301 GTGGGAGAGTCTGTAAGGCTGATTCAGGGGAGAGCAGCCCTGACAGAAATGAGAC 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 994 ACGTTCCTGAGAGCGGCGGTGACATGCTGTGCCCCGAACTGGCAGCTACTTCAACGAG 1053  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 361 ACGTTCCTGAGAGCGGCGGTGACATGCTGTGCCCCGAACTGGCAGCTACTTCAACGAG 420  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1054 CTGACGGGAGTACACCCCTGACAGACCTCTCTGACACACACCTCTTTGGGGT 1113  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 421 CTGAAGGCGGTGACACCCCTGAGGGGCGCTCTTGGACACACACCTTTCTGGGGT 480  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1114 TTTCAGACCACTGAGGCTGACATGTAACCTGTGCGCATCTGTAGTACAGTTGGAAGAG 1173  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 481 TTTCAGGCGCATGAGGGGCGATATGACCTGTGAGTTGTGAGTACAGTTGGAAGAG 540  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1174 ATCCAGGCGGTGTTGAGAGGCGCCCTATTAAGAGTACATGAGGAAGCCAGAAAGTGAGC 1233  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 541 ATCCAGCAAGTGTGAGAGGCTCCCTACAAAGAGTACAGTACAGCCAGAAAGTGAGCC 600  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1234 CGCTACACTGACCTGTATACCCAGCCCTGCGCTGCTGTGATTAACAACTGGCATGG 1293  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 601 CCTTACTGACCTGACCTGACCCCTGAGCTGCTGTATCAACAACTGGCAACGA 660  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1294 CGCCAGCGGTACACAGCTCCCTGAGAGTACCCGCAACATCTCAACTCTGCAAGAAG 1353  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 661 GCAATGCTTACACAGTTCCTGTGAGAGTCCCGGAGC-ACACCTCAACTCATCAAGAAG 719  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 1354 CACCC 1358  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 720 CACCC 724

RESULT 15  
 LOCUS BG324157 597 bp mRNA linear EST 27-FEB-2001  
 DEFINITION BG324157 NIH-MGC\_14 Homo sapiens cDNA clone IMAGE:4561275 5',  
 mRNA sequence.  
 BG324157  
 ACCESSION BG324157.1 GI:13130594  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE Unpublished  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bhs-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at:  
 http://image.llnl.gov  
 Plate: L12M1271 row: k column: 04  
 High quality sequence stop: 594.  
 Location/Qualifiers  
 1..597  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4561275"  
 /tissue\_type="renal cell adenocarcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: kidney; Vector: pOTB; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).\*

BASE COUNT 130 a 176 c 175 g 116 t  
ORIGIN

Query Match 23.5%; Score 536.4; DB 10; Length 597;

Best Local Similarity 98.3%; Pred. No. 3.3e-113; Mismatches 6; Indels 4; Gaps 4;

Matches 584; Conservative 0; Mismatches 6; Indels 4; Gaps 4;

```
QY 1034 GGCAGCTCTACTTCAACACGCTGCAGGCGATGCACACCCCTGCAGACACTCTGGCACA 1093
DB 2 GGCAGCTCTACTTCAACACGCTGCAGGCGATGCACACCCCTGCAGACACTCTGGCACA 61
QY 1094 ACACACCTCTTTGGGGTTTTCAGCAGAGTGGGGTGACATGTACTGTGGCCATCT 1153
DB 62 ACACACCTCTT-TTGGGGTTTATCAGCAGAGTGGGGTGACATGTACTGTGGCCATCT 120
QY 1154 GTGAGTACCATTTGGAAGATCCAGGGGGTGTGAGGGGCCCTATAGGAGTACCATG 1213
DB 121 GTGAGTACCAT-TGGAAGATCCAGGGGGTGTGAGGGGCCCTATAGGAGTACCATG 179
QY 1214 AGGAAGCCCAAGATGGGACCGCTACACTGACCCCTGTACCCAGCCCTGGCTGCTGT 1273
DB 180 AGGAAGCCCAAGATGGGACCGCTACACTGACCCCTGTACCCAGCCCTGGCTGCTGT 239
QY 1274 GCATTACACTGGCAGTGGGCCACGGCTACACCAAGCTCCCTGGAGCTACCCGACA 1333
DB 240 GCATTACACTGGCAGTGGGCCACGGCTACACCAAGCTCCCTGGAGCTACCCGACA 299
QY 1334 TCCTCAACTTGTCAAGAGACACCGGCTGATGAGAGGACAGTGGGGCTCGGTGGAGCC 1393
DB 300 TCCTCAACTTGTCAAGAGACACCGGCTGATGAGAGGACAGTGGGGCTCGGTGGAGCC 359
QY 1394 GCCCCTGCTGCTGTAAGAGGAGGACCACTTACACCACCTGCTGGCCGACCGGTTACAG 1453
DB 360 TGCCCTGCTGCTGTAAGAGGAGGACCACTTACACCACCTGCTGGCCGACCGGTTACAG 419
QY 1454 GACTTGATGAGACCACTATACAGTGTGTTGATGAGCAGAGAGCGGCTGG-CTGCTC 1512
DB 420 GACTTGATGAGACCACTATACAGTGTGTTGATGAGCAGAGAGCGGCTGGACTGCTC 479
QY 1513 AAGGCTGTGAGCCTGGGGCCCTGGGGTTCACCTGATTGAGAGCTGCAGCTG-TTGGACCA 1571
DB 480 AAGGCTGTGAGCCTGGGGCCCTGGGGTTCACCTGATTGAGAGCTGCAGCTGTTGGACCA 539
QY 1572 GGAGCCCATGAGAGCCTGTGTATCTCAGAGCAAGAGCTGCTTTGCCGG 1625
DB 540 GGAGCCCATGAGAGCCTGTGTATCTCAGAGCAAGAGCTGCTTTTGCGG 593
```

Search completed: August 19, 2003, 15:52:41  
Job time : 3213 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 12:26:06 ; Search time 437 Seconds  
(without alignments)  
14108.733 Million cell updates/sec

Title: us-10-002-050-13

Perfect score: 2284  
Sequence: 1 cggcccttcacacacccctcgtg.....tgtgtgtgtgtatgtgtgtgt 2284

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2280	99.8	2284	22	AAC84888 Human SEC7 nucleic
2	1887.4	82.6	2155	22	AAC84887 Human SEC6 nucleic
3	1879.4	82.3	2156	22	AAC84892 Human SEC11 nucleic
4	1761	77.1	3293	22	AAD08048 Human extracellular
5	1738	76.1	3868	24	AAD28947 Human MOL5a cDNA
6	1722	75.4	3112	24	AAD28949 Human MOL5c cDNA
7	1699	74.4	2558	24	AAD28948 Human MOL5b cDNA
8	1617.6	70.8	3776	24	ABN63963 Human gene sequenc

9	1394	61.0	8095	24	ABX92031 Lung specific nucl
10	1394	61.0	8144	25	ABT13397 Breast specific re
11	1351	59.2	3503	22	AAE29461 Murine M-Sema-F CD
12	1193.6	52.3	2811	22	AAE29451 Human TRNGO 276 CD
13	1094.2	47.9	2813	24	ABN59626 Novel human cooling
14	928.4	40.6	1024	24	ABN85379 Partial human NOV2
15	841.4	36.8	967	22	AAE30193 Clone 1498905.0.6
16	463	20.3	729	22	AAE29452 Human TRNGO 276 CD
17	423	18.5	2132	24	ABL50334 Human cancer cell
18	411	18.0	3781	21	AA37092 Human PRO1480 (UNO
19	411	18.0	3781	22	AA56151 Human DNA encoding
20	411	18.0	3781	22	AAE54381 Primer #75 used in
21	411	18.0	3781	25	Human PRO1480 cDNA
22	411	18.0	3781	25	Human cDNA encodin
23	411	18.0	3781	25	Human cDNA encodin
24	411	18.0	3781	25	Novel human secret
25	411	18.0	3781	25	Human secreted/tra
26	411	18.0	3781	25	Human PRO polynuc
27	411	18.0	3781	25	Human PRO polynuc
28	411	18.0	3781	25	Human cDNA encodin
29	411	18.0	3781	25	Human PRO polynuc
30	409.4	17.9	3556	22	Human cDNA encodin
31	409.4	17.9	3766	24	ABX77938 Hypoxia induced pr
32	409.4	17.9	3766	24	ABN83984 Human gene sequenc
33	409.4	17.9	3766	25	ACC51059 Human bladder canc
34	409.4	17.9	3766	25	ABX76369 Lung cancer-assoc
35	409.4	17.9	2703	21	AAA47437 Sequence encoding
36	406.8	17.8	2390	24	ABR11109 DNA encoding human
37	401	17.6	4157	18	AAE50655 Human CD100 antige
38	401	17.6	4157	22	AAE77463 Human CD100 DNA se
39	401	17.6	4157	22	AAE77463 Human CD100 nucle
40	398.2	17.4	2769	19	AAV31121 Human semaphorin e
41	398.2	17.4	2769	22	AAE77462 Murine CD100 DNA s
42	398.2	17.4	2769	22	AAE77462 Mouse CD100 nucle
43	398.2	17.4	4391	18	AAE50656 Mouse CD100 antige
44	370.4	16.2	2517	24	ABR4076 Human semaphorin G
45	370.4	16.2	2598	24	ABA04078 Human semaphorin G

## ALIGNMENTS

RESULT 1	
AAC84888	
ID	AAC84888 standard; cDNA; 2284 BP.
XX	
AC	AAC84888;
XX	
DT	20-APR-2001 (first entry)
XX	
DE	Human SEC7 nucleic acid sequence (clone ID 20422974.2).
XX	
KW	SECX; cytosolic; gynecological; gene therapy; screening assay; human;
KW	chromosomal mapping; forensic biology; cell proliferation; cancer;
KW	cell differentiation; immune associated disorder; gestational disease;
KW	SEC6; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "SEC7"
XX	
PN	WO200078802-A2.
XX	
PD	28-DEC-2000.
XX	
PF	23-JUN-2000; 2000WO-US17328.
XX	
PR	23-JUN-1999; 99US-0140584.
PR	20-JUL-1999; 99US-0144722.
PR	16-SEP-1999; 99US-0154520.

PR 22-JUN-2000; 2000US-0604286.  
 XX (CURA-) CURAGEN CORP.  
 XX Shinkens RA, Fernandes E, Vernet C, Yang M, Boldog FL;  
 PI Herrmann JL;  
 XX WPI; 2001-071385/08.  
 DR P-PSDB; AAB48374.  
 XX polynucleotides encoding SECX proteins useful for treating disease  
 PT characterized by an aberrant level of cell proliferation and/or  
 PT differentiation like cancer or immune associated disorders -  
 XX  
 PS Claim 3; Fig 7; 132p; English.  
 XX  
 CC The invention relates to human SECX polypeptides and polynucleotides  
 CC encoding them. The SECX polypeptides can be expressed by standard  
 CC recombinant methodology. The SECX polypeptides are useful for treating  
 CC or preventing a SECX-associated disorder. The invention is useful in  
 CC screening assays; detection assays (e.g. chromosomal mapping, cell and  
 CC tissue typing, forensic biology); predictive medicine (diagnostic assays,  
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and  
 CC methods of treatment (e.g. therapeutic and prophylactic), especially  
 CC disorders characterized by aberrant cell proliferation and/or  
 CC differentiation like cancer or immune associated disorders or gestational  
 CC disease. The present sequence represents a SEC7 nucleic acid sequence.  
 CC  
 XX  
 SQ Sequence 2284 BP; 443 A; 683 C; 689 G; 465 T; 4 other;  
 Query Match 99.8%; Score 2280; DB 22; Length 2284;  
 Best Local Similarity 100.0%; Freq. No. 0;  
 Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGCCCTTCTACACCTCTGCTGCTGATGTGAAGCGGGTTGGGGTTCTGAGAGCT 60  
 1 CGGCCCTTCTACACCTCTGCTGCTGATGTGAAGCGGGTTGGGGTTCTGAGAGCT 60  
 61 ATTGTCGCGCTGGGGAAGGAGGAGCGCGGACCTCCGCTGCGACCGCGCCG 120  
 61 ATTGTCGCGCTGGGGAAGGAGGAGCGCGGACCTCCGCTGCGACCGCGCCG 120  
 61 ATTGTCGCGCTGGGGAAGGAGGAGCGCGGACCTCCGCTGCGACCGCGCCG 120  
 121 ACCACAGAGACGCTGGGCTGAAAGCTCAGAGCGGGGGCTGGCCATGGCCCACTAG 180  
 121 ACCACAGAGACGCTGGGCTGAAAGCTCAGAGCGGGGGCTGGCCATGGCCCACTAG 180  
 121 ACCACAGAGACGCTGGGCTGAAAGCTCAGAGCGGGGGCTGGCCATGGCCCACTAG 180  
 181 GCTGTCGCTGCTGCTGAGACAGCTGTGGGCTGCGCATGGGGCTGAGGTGTGTTG 240  
 181 GCTGTCGCTGCTGCTGAGACAGCTGTGGGCTGCGCATGGGGCTGAGGTGTGTTG 240  
 181 GCTGTCGCTGCTGCTGAGACAGCTGTGGGCTGCGCATGGGGCTGAGGTGTGTTG 240  
 241 AACCTTGTCCGCGTAAGACAGTGTCTCTGCGGAGCTGGCCACGGTAGTACGGCGTTT 300  
 241 AACCTTGTCCGCGTAAGACAGTGTCTCTGCGGAGCTGGCCACGGTAGTACGGCGTTT 300  
 301 TCCGAGACCGGATCCAGGACTTCTGACACTGAGAGCTGAGAGCCCACTGGGCTTTG 360  
 301 TCCGAGACCGGATCCAGGACTTCTGACACTGAGAGCTGAGAGCCCACTGGGCTTTG 360  
 301 TCCGAGACCGGATCCAGGACTTCTGACACTGAGAGCTGAGAGCCCACTGGGCTTTG 360  
 361 TACGTGGGCGCCGAGAGAGCCCTGTTTGCCTTACAGATGAGAGCCCTGAGGTGCAAGA 420  
 361 TACGTGGGCGCCGAGAGAGCCCTGTTTGCCTTACAGATGAGAGCCCTGAGGTGCAAGA 420  
 361 TACGTGGGCGCCGAGAGAGCCCTGTTTGCCTTACAGATGAGAGCCCTGAGGTGCAAGA 420  
 421 GCGATCTCTGGGAGAGCCCGCTGGAGAGAAAGACTGAGTGTATCCAGAAAGGAAGAC 480  
 421 GCGATCTCTGGGAGAGCCCGCTGGAGAGAAAGACTGAGTGTATCCAGAAAGGAAGAC 480  
 421 GCGATCTCTGGGAGAGCCCGCTGGAGAGAAAGACTGAGTGTATCCAGAAAGGAAGAC 480  
 481 AACCGAGCGAGTGTTCATCTTCACTTCCGCTTCCGAGAGCCCTTAAATGCGCTCCACCTG 540  
 481 AACCGAGCGAGTGTTCATCTTCACTTCCGCTTCCGAGAGCCCTTAAATGCGCTCCACCTG 540  
 541 TACGCTGTGGACACTACGCTTCCAGGCCCAAGTACACTACGTCAACATGCTCACTTC 600  
 541 TACGCTGTGGACACTACGCTTCCAGGCCCAAGTACACTACGTCAACATGCTCACTTC 600

QY 601 ACTTTGAGCAGTGAAGAGTTTGAAGTGGAAAGGCAAGTGTCCCTATGACCCAGTAAG 660  
 DB 601 ACTTTGAGCAGTGAAGAGTTTGAAGTGGAAAGGCAAGTGTCCCTATGACCCAGTAAG 660  
 QY 661 GGCATGCTGCGCTTCTTGTGATGAGTGTAGTGTACTGCGCCCACTCAACAACTTCTG 720  
 DB 661 GGCATGCTGCGCTTCTTGTGATGAGTGTAGTGTACTGCGCCCACTCAACAACTTCTG 720  
 QY 721 GGCAGGAAAGCAATATCTGCTGAACATGAGGGGCCCCACACTCCATGAAGACAGTAC 780  
 DB 721 GGCAGGAAAGCAATATCTGCTGAACATGAGGGGCCCCACACTCCATGAAGACAGTAC 780  
 QY 781 CTGGCTTTTGGCTCAACGACCTCACTTGTAGGCTGTGCTATGTACTAGAGTGTG 840  
 DB 781 CTGGCTTTTGGCTCAACGACCTCACTTGTAGGCTGTGCTATGTACTAGAGTGTG 840  
 QY 841 GGCAGCTTACGAGGAGGAGCAGACAAAGTCTACTTCTTCAAGGAGCGGCGATGAG 900  
 DB 841 GGCAGCTTACGAGGAGGAGCAGACAAAGTCTACTTCTTCAAGGAGCGGCGATGAG 900  
 QY 901 TCCGACTGTATGCGAGGAGGAGTGTGCTGCTGAGGCGGCTGCAAGGCGATATG 960  
 DB 901 TCCGACTGTATGCGAGGAGGAGTGTGCTGCTGAGGCGGCTGCAAGGCGATATG 960  
 QY 961 GGGGGGCGACGAGACCTGACAGAGAAAGTGAACACGTTCTTCAAGGCGGCGATGAC 1020  
 DB 961 GGGGGGCGACGAGACCTGACAGAGAAAGTGAACACGTTCTTCAAGGCGGCGATGAC 1020  
 QY 1021 TCTGCCCCGAATGAGAGTCTACTTCAACAGCTGACAGGAGTACACCTGACAGAC 1080  
 DB 1021 TCTGCCCCGAATGAGAGTCTACTTCAACAGCTGACAGGAGTACACCTGACAGAC 1080  
 QY 1081 ACCCTGCGACAAACACACCTTCTTGGGGTTTTCAGACAGTGGGGTACATGTAC 1140  
 DB 1081 ACCCTGCGACAAACACACCTTCTTGGGGTTTTCAGACAGTGGGGTACATGTAC 1140  
 QY 1141 CTGTGGCCATCTGTGATGACAGTGTGAAGAGATTCACAGGGGTGTTGAGGGCCCTAT 1200  
 DB 1141 CTGTGGCCATCTGTGATGACAGTGTGAAGAGATTCACAGGGGTGTTGAGGGCCCTAT 1200  
 QY 1201 AAGGATACATGAGGAAGCCCAAGAGTGGAGCCGTACACTGACCTTACCCAGCCT 1260  
 DB 1201 AAGGATACATGAGGAAGCCCAAGAGTGGAGCCGTACACTGACCTTACCCAGCCT 1260  
 QY 1261 CGGCTGCGCTGCTGATTAACAAGTGGCATCGGGCGCACAGCTACACAGCTCCCTGGAG 1320  
 DB 1261 CGGCTGCGCTGCTGATTAACAAGTGGCATCGGGCGCACAGCTACACAGCTCCCTGGAG 1320  
 QY 1321 CTACCCGACACATCCTCAACTTCTGTAAGAAAGACCCGCTATGAGAGAGAGTGGGG 1380  
 DB 1321 CTACCCGACACATCCTCAACTTCTGTAAGAAAGACCCGCTATGAGAGAGAGTGGGG 1380  
 QY 1381 CCTGGTGAAGGCGCCCTCTGCTGTAAGAAAGGACCAACTTCAACCACTGCTGGGC 1440  
 DB 1381 CCTGGTGAAGGCGCCCTCTGCTGTAAGAAAGGACCAACTTCAACCACTGCTGGGC 1440  
 QY 1441 GACCGGGTTACAGACTTGTATGAGACCACTTATACAGTCTGTTTCAATGGACAGAGAC 1500  
 DB 1441 GACCGGGTTACAGACTTGTATGAGACCACTTATACAGTCTGTTTCAATGGACAGAGAC 1500  
 QY 1501 GGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGAGTTCACTGATGAGAGCTGACG 1560  
 DB 1501 GGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGAGTTCACTGATGAGAGCTGACG 1560  
 QY 1561 CTGTTTGAACAGAGACCCATGAGAAAGCTGTGTCTATCTCAAGAGAAAGTGTCTTT 1620  
 DB 1561 CTGTTTGAACAGAGACCCATGAGAAAGCTGTGTCTATCTCAAGAGAAAGTGTCTTT 1620  
 QY 1621 GCGGCTCCCGCTCAGAGTGTGACAGTGGCCGAGGCGGACGTGCAATGAAGTATGCTTC 1680  
 DB 1621 GCGGCTCCCGCTCAGAGTGTGACAGTGGCCGAGGCGGACGTGCAATGAAGTATGCTTC 1680

[illegible]

PR	23-JUN-1999;	99US-0140584.
PR	20-JUL-1999;	99US-0144722.
PR	16-SEP-1999;	99US-0154520.
PR	22-JUN-2000;	2000US-0604286.
XX		
PA	(CURA-) CURAGEN CORP.	
XX		
PI	Shimkets RA, Fernandes E, Vernet C, Yang M, Boldog FI.	
PI	Herrmann JL;	
DR	WPI: 2001-071385/08.	
XX	P-PSDB; ABA48373.	
PT	Polynucleotides encoding SECX proteins useful for treating disease	
PT	characterized by an aberrant level of cell proliferation and/or	
PT	differentiation like cancer or immune associated disorders	
XX		
PS	Claim 3; Fig 6; 132pp; English.	
XX		
CC	The invention relates to human SECX polypeptides and polynucleotides	
CC	encoding them. The SECX polypeptides can be expressed by standard	
CC	recombinant methodology. The SECX polypeptides are useful for treating	
CC	or preventing a SECX-associated disorder. The invention is useful in	
CC	screening assays; detection assays (e.g. chromosomal mapping, cell and	
CC	tissue typing, forensic biology); predictive medicine (diagnostic assays, and	
CC	prognostic assays, monitoring clinical trials, and pharmacogenomics); and	
CC	methods of treatment (e.g. therapeutic and prophylactic), especially	
CC	disorders characterized by aberrant cell proliferation and/or	
CC	differentiation like cancer or immune associated disorders or gestational	
CC	disease. The present sequence represents a SEC6 nucleic acid sequence.	
XX		
SO	Sequence 2155 BP; 444 A; 614 C; 645 G; 452 T; 0 other;	
	Query Match 82.6%; Score 1887.4; DB 22; Length 2155;	
	Best Local Similarity 99.5%; Pred. No. 0;	
	Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2	
OY	1	CGGCCCTTCTACACTCTCGCCCTGCTGATGGAACGGGGTTGGGGTTTCGACAGGCT 60
DB	1	CGGCCCTTCTACACTCTCGCCCTGCTGATGGAACGGGGTTGGGGTTTCGACAGGCT 60
OY	61	ATTGTCTGCGTGGGGAAGGGGACAGCGCCGGGACCTCGCTCGACAGCGCGCG 120
DB	61	ATTGTCTGCGTGGGGAAGGGGACAGCGCCGGGACCTCGCTCGACAGCGCGCG 120
OY	121	ACCGACAGACAGACGCGCTGAAGCTAGAGCGCGGGGGCTGGGCCATGGCCCCACACTGG 180
DB	121	ACCGACAGACAGACGCGCTGAAGCTAGAGCGCGGGGGCTGGGCCATGGCCCCACACTGG 180
OY	181	GCTGCTGCGTGCAGACAAAGCTGTGGGGCTGGGCATTTGGGGCTGAGGTGTGTGTGG 240
DB	181	GCTGCTGCGTGCAGACAAAGCTGTGGGGCTGGGCATTTGGGGCTGAGGTGTGTGTGG 240
OY	241	AACCTTGTGCGCGCTTAAGACAGTGTCTCTGTGGGAGCTGGCCACGCTAGTACGGCGTTTC 300
DB	241	AACCTTGTGCGCGCTTAAGACAGTGTCTCTGTGGGAGCTGGCCACGCTAGTACGGCGTTTC 300
OY	301	TCCACAGACCGGCATTCAGAGACTTCTGTACACTGAGCGTGTAGCGAGCCCTGGCTTCTG 360
DB	301	TCCACAGACCGGCATTCAGAGACTTCTGTACACTGAGCGTGTAGCGAGCCCTGGCTTCTG 360
OY	361	TACGTTGCGCGCGCGAGAGCCCTGTTTGCCTTCACAGTGTAGAGCCCTGAGAGTGCAGGA 420
DB	361	TACGTTGCGCGCGCGAGAGCCCTGTTTGCCTTCACAGTGTAGAGCCCTGAGAGTGCAGGA 420
OY	421	GCGATCTCTGTGGAGGCCCGGTGAGAAAGACTGATGTATCCAGAAAGGGAAGAAC 480
DB	421	GCGATCTCTGTGGAGGCCCGGTGAGAAAGACTGATGTATCCAGAAAGGGAAGAAC 480
OY	481	AACGACAGCAGGTGTTCACTTATATCGCTTCGCGAGCCCTACAAATGCCCTCCACCTG 540
DB	481	AACGACAGCAGGTGTTCACTTATATCGCTTCGCGAGCCCTACAAATGCCCTCCACCTG 540

541 TACGCTGTGGACCTAGACCTTCCAGCCCAAGTGCACCTACGTCAACATGCTCACCTTC 600  
 541 TACGCTGTGGACCTAGACCTTCCAGCCCAAGTGCACCTACGTCAACATGCTCACCTTC 600  
 601 ACTTGGAGCATGAGAGTTTGAAGATGGAGAGGCAAGTGTCCCTATGACCACCTAAG 660  
 601 ACTTGGAGCATGAGAGTTTGAAGATGGAGAGGCAAGTGTCCCTATGACCACCTAAG 660  
 661 GGCATGTCTGGCTTCTGTGTGATGAGTGTACTGTGGCCACACCTCAACAACTTCTG 720  
 661 GGCATGTCTGGCTTCTGTGTGATGAGTGTACTGTGGCCACACCTCAACAACTTCTG 720  
 721 GGCAGGAAACCATATATCTGTGTGATGAGTGTACTGTGGCCACACCTCAACAACTTCTG 780  
 721 GGCAGGAAACCATATATCTGTGTGATGAGTGTACTGTGGCCACACCTCAACAACTTCTG 780  
 781 CTGGCTTTTGGCTCAAGCAACCTCATTGTAGCTGTGCTATGATGATGAGTGTG 840  
 781 CTGGCTTTTGGCTCAAGCAACCTCATTGTAGCTGTGCTATGATGATGAGTGTG 840  
 841 GGCAGCTTACGGGGGAGCAGACAGCAAGGTCTACTTCTTTCAGGAGCGGCGAGTGGAG 900  
 841 GGCAGCTTACGGGGGAGCAGACAGCAAGGTCTACTTCTTTCAGGAGCGGCGAGTGGAG 900  
 901 TCCGACTGTATGCGAGAGGTGGTGGCTGTGTGGCCCTGTGCAAGGGCGATATG 960  
 901 TCCGACTGTATGCGAGAGGTGGTGGCTGTGTGGCCCTGTGCAAGGGCGATATG 960  
 961 GGGGGCGACGAGACCTGTGAGAGAGTGCACAGCTTCTGTAAGCGCGCTGGCATGC 1020  
 961 GGGGGCGACGAGACCTGTGAGAGAGTGCACAGCTTCTGTAAGCGCGCTGGCATGC 1020  
 1021 TCTGCCCCCACTGTCAGTCTACTTCAACAGCTGACAGCGCATGACACCTGAGGAC 1080  
 1021 TCTGCCCCCACTGTCAGTCTACTTCAACAGCTGACAGCGCATGACACCTGAGGAC 1080  
 1081 ACCTCTGGCAACACCACTCTTTGGGGTTTTCAGGACAGTGGGGTACATGTAC 1140  
 1081 ACCTCTGGCAACACCACTCTTTGGGGTTTTCAGGACAGTGGGGTACATGTAC 1140  
 1141 CTGTGCGGCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 1141 CTGTGCGGCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 1201 AAGGAGTACATGAGAGAGCCAGAGAGTGGACCCCTACACTGACCTGTACCCAGCCCT 1260  
 1201 AAGGAGTACATGAGAGAGCCAGAGAGTGGACCCCTACACTGACCTGTACCCAGCCCT 1260  
 1261 GGGCTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 1261 GGGCTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 1321 CTACCCGCAACATCTCAACTTCTCAAGAGACCCCGCTGATGAGAGAGTGGAG 1380  
 1321 CTACCCGCAACATCTCAACTTCTCAAGAGACCCCGCTGATGAGAGAGTGGAG 1380  
 1381 CCTGCTGGAGAGCCCGCTGCTGTGAAGAGGACCACTTACCCACCTGTGTGCC 1440  
 1381 CCTGCTGGAGAGCCCGCTGCTGTGAAGAGGACCACTTACCCACCTGTGTGCC 1440  
 1441 GACCGGTTTACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 1441 GACCGGTTTACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 1501 GGGCTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 1501 GGGCTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 1561 CTGTTTTCAGAGAGCCCATGAGAGCCCTGTGTGATGATGATGATGATGATGATGAT 1620  
 1561 CTGTTTTCAGAGAGCCCATGAGAGCCCTGTGTGATGATGATGATGATGATGATGAT 1620  
 1621 GCCGGCTCCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

1621 GCGGGCTCCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 1681 TGTGAGATGTGTCTGTGAG 1740  
 1681 TGTGAGATGTGTCTGTGAG 1740  
 1741 TGTGAG 1800  
 1741 TGTGAG 1800  
 1801 ACTTACAGGATTTGCAACCTCCGTGAGATGATGATGATGATGATGATGATGATGATGAT 1860  
 1801 ACTTACAGGATTTGCAACCTCCGTGAGATGATGATGATGATGATGATGATGATGATGAT 1860  
 1861 ATCAGGATGTGAG 1920  
 1861 ATCAGGATGTGAG 1920  
 1921 CTG 1923  
 1918 CCG 1920  
 1918 CCG 1920  
 RESULT 3  
 AAC84892  
 ID AAC84892 standard; cDNA; 2156 BP.  
 AAC84892;  
 20-APR-2001 (first entry)  
 Human SEC11 nucleic acid sequence (clone ID 20422974.0.132-ext2).  
 DE  
 SE: cytoskeletal; gynecological; gene therapy; screening assay; human;  
 KW chromosomal mapping; forensic biology; cell proliferation; cancer;  
 KW cell differentiation; immune associated disorder; gestational disease;  
 SEC11; 88.  
 OS  
 Homo sapiens.  
 OS  
 Key Location/Qualifiers  
 CDS 166..2040  
 FT /tag= a  
 FT /product= "SEC11"  
 W020078802-A2.  
 28-DEC-2000.  
 23-JUN-2000; 2000MO-US17328.  
 23-JUN-1999; 99US-0140584.  
 20-JUL-1999; 99US-0144722.  
 16-SEP-1999; 99US-0154520.  
 22-JUN-2000; 2000US-0604286.  
 (CURA-) CURAGEN CORP.  
 Shinkels RA, Fernandes E, Vernet C, Yang M, Boldog FL,  
 Herrmann JL,  
 WPI; 2001-071385/08.  
 P-PSDB; AAB48378.  
 polynucleotides encoding SECX proteins useful for treating disease  
 characterized by an aberrant level of cell proliferation and/or  
 differentiation like cancer or immune associated disorders -  
 Claim 3; Fig 11; 132pp; English.  
 The invention relates to human SECX polypeptides and polynucleotides  
 encoding them. The SECX polypeptides can be expressed by standard



CC recombinant methodology. The SECK polypeptides are useful for treating  
CC or preventing a SECK-associated disorder. The invention is useful in  
CC screening assays; detection assays (e.g. chromosomal mapping, cell and  
tissue typing, forensic biology); predictive medicine (diagnostic assays,  
CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and  
CC methods of treatment (e.g. therapeutic and prophylactic), especially  
CC disorders characterized by aberrant cell proliferation and/or  
CC differentiation like cancer or immune associated disorders or gestational  
CC disease. The present sequence represents a SECK1 nucleic acid sequence.

XX Sequence 2156 BP; 444 A; 617 C; 645 G; 450 T; 0 other;

Query Match 82.3%; Score 1879.4; DB 22; Length 2156;

Best Local Similarity 99.3%; Pred. NO. 0;  
Matches 1909; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 1 CGGCCCTTCTCACACCTCCCTGCTGATGTGAACGGGGTTGGGGTCTGAGGGCT 60  
DB 1 CGGCCCTTCTCACACCTCCCTGCTGATGTGAACGGGGTTGGGGTCTGAGGGCT 60  
QY 61 ATTGTCTGCGTGGGAAAGGGGACAGGCGGAGACCGGAGCTCGCTCGGAGCGGCGC 120  
DB 61 ATTGTCTGCGTGGGAAAGGGGACAGGCGGAGACCGGAGCTCGCTCGGAGCGGCGC 120  
QY 121 ACCAGCAGACAGCTGGCTGAAAGCTCAGAGCCGGGGCTGCGCCATGCCCCACACTGG 180  
DB 121 ACCAGCAGACAGCTGGCTGAAAGCTCAGAGCCGGGGCTGCGCCATGCCCCACACTGG 180  
QY 181 GCGTCTGCGTGGTGGAGCAAGCTGTGGGGCTGGGGCATTTGGGGTGGAGTGTGGTGG 240  
DB 181 GCGTCTGCGTGGTGGAGCAAGCTGTGGGGCTGGGGCATTTGGGGTGGAGTGTGGTGG 240  
QY 241 AACCTTGTGCGGCGTGAAGAGTGTCTTGGGGAGCTGGGCAAGTGTGAGGCGGCTTC 300  
DB 241 AACCTTGTGCGGCGTGAAGAGTGTCTTGGGGAGCTGGGCAAGTGTGAGGCGGCTTC 300  
QY 301 TCCCAAGCCGGCATCCAGACTTCTGACACTGACCTGACAGGAGCCACTGGGCTTCTG 360  
DB 301 TCCCAAGCCGGCATCCAGACTTCTGACACTGACCTGACAGGAGCCACTGGGCTTCTG 360  
QY 361 TAGCTGGGCGCCGAGAGGCGCTGTTGCTTCGATGAGTGGAGGCGCTGGAGCTCAAGGA 420  
DB 361 TAGCTGGGCGCCGAGAGGCGCTGTTGCTTCGATGAGTGGAGGCGCTGGAGCTCAAGGA 420  
QY 421 GCGATCTCTGGGAGGCGCCCGTGGAGAGAAAGTGTATCCAGAAAGGAGAAAC 480  
DB 421 GCGATCTCTGGGAGGCGCCCGTGGAGAGAAAGTGTATCCAGAAAGGAGAAAC 480  
QY 481 AACCAAGCAGAGTGTCAACTTCATCCGCTTCTGAGCCCTACATGCTCCCACTG 540  
DB 481 AACCAAGCAGAGTGTCAACTTCATCCGCTTCTGAGCCCTACATGCTCCCACTG 540  
QY 541 TAGCTGTGGCACTACGCTTCCACGCAAGTGAACCTACGTAACATGCTTCACTTC 600  
DB 541 TAGCTGTGGCACTACGCTTCCACGCAAGTGAACCTACGTAACATGCTTCACTTC 600  
QY 601 ACTTTGGAGCATGAGAGTGTGAAGATGGGAGGGAAGTGTCCCTATGACCCAGTAAG 660  
DB 601 ACTTTGGAGCATGAGAGTGTGAAGATGGGAGGGAAGTGTCCCTATGACCCAGTAAG 660  
QY 661 GGGCATGCTGGCTTCTTGTGATGTGTAGCTGTACTCGGCACTCAACACTTCTG 720  
DB 661 GGGCATGCTGGCTTCTTGTGATGTGTAGCTGTACTCGGCACTCAACACTTCTG 720  
QY 721 GGGCAGCAACCAATATCTGCGCTTAACATGGGGGCCCACTCAATGAAGAGAGTAC 780  
DB 721 GGGCAGCAACCAATATCTGCGCTTAACATGGGGGCCCACTCAATGAAGAGAGTAC 780  
QY 781 CTGGCCTTTTGGCTCAAGAACTCACTTTAGGCTGTGCTGCTATGACCTGAAGTGTG 840  
DB 781 CTGGCCTTTTGGCTCAAGAACTCACTTTAGGCTGTGCTGCTATGACCTGAAGTGTG 840  
QY 841 GGCAGCTTCAAGGGGAGCAGACAGGTCTACTTCTTTCAGGGAGCGGGCAGTGGAG 900

DB 841 GGCAGCTTCAAGGGGAGCAGACAGGTCTACTTCTTTCAGGGAGCGGGCAGTGGAG 900  
QY 901 TCCGACTGATATGCGGAGAGTGTGGCTCGTGTGGCCGCTGTCTGCAAGGGCGATATG 960  
DB 901 TCCGACTGATATGCGGAGAGTGTGGCTCGTGTGGCCGCTGTCTGCAAGGGCGATATG 960  
QY 961 GGGGGCGCAGGACCTGCAAGAGAAAGTGAACCACTTCTGAAAGCGCGCTGCGATGC 1020  
DB 961 GGGGGCGCAGGACCTGCAAGAGAAAGTGAACCACTTCTGAAAGCGCGCTGCGATGC 1020  
QY 1021 TGTGCCCCGAATGCGAGCTCTACTTCAACAGCTGAGCGGATGCAACCTGCAAGAC 1080  
DB 1021 TGTGCCCCGAATGCGAGCTCTACTTCAACAGCTGAGCGGATGCAACCTGCAAGAC 1080  
QY 1081 ACCCTCTGGCAACACACCTTCTTGGGGTTTTCACACAGTGGGGTGCATGTAC 1140  
DB 1081 ACCCTCTGGCAACACACCTTCTTGGGGTTTTCACACAGTGGGGTGCATGTAC 1140  
QY 1141 CTGTGCGCATCTGTGAGTACAGATTGGAAGACATCCAGCGGGTGTGAGGGCCCTAT 1200  
DB 1141 CTGTGCGCATCTGTGAGTACAGATTGGAAGACATCCAGCGGGTGTGAGGGCCCTAT 1200  
QY 1201 AAGGATACCATGAGAAAGCCCAAGAGTGGAGCCGTACACTGACCTGTACCAAGCCT 1260  
DB 1201 AAGGATACCATGAGAAAGCCCAAGAGTGGAGCCGTACACTGACCTGTACCAAGCCT 1260  
QY 1261 CGGCTGGGCTGTGATTAACAATGAGCATGCGGCGCAGCGCTACACCACTCCCTGAG 1320  
DB 1261 CGGCTGGGCTGTGATTAACAATGAGCATGCGGCGCAGCGCTACACCACTCCCTGAG 1320  
QY 1321 CTACCCGCAACAATCTCAACTTCGTCAAGAAGCACCCTGTGATGAGAGCAGAGTGGG 1380  
DB 1321 CTACCCGCAACAATCTCAACTTCGTCAAGAAGCACCCTGTGATGAGAGCAGAGTGGG 1380  
QY 1381 CTTGCGTGGAGCGCCCTCTGCTGTGAGAAAGGGCACCAATTCACCCACTGTGGCC 1440  
DB 1381 CTTGCGTGGAGCGCCCTCTGCTGTGAGAAAGGGCACCAATTCACCCACTGTGGCC 1440  
QY 1441 GACCGGGTTACAGACTTGTGAGAGCCACCTATACATGCTGTTTATGGAGCAGAGAC 1500  
DB 1441 GACCGGGTTACAGACTTGTGAGAGCCACCTATACATGCTGTTTATGGAGCAGAGAC 1500  
QY 1501 GCGTGGCTGTCAAGGCTGTGAGGCTGTGGGCGCTGAGTCACTGATTTGAGAGCTGAG 1560  
DB 1501 GCGTGGCTGTCAAGGCTGTGAGGCTGTGGGCGCTGAGTCACTGATTTGAGAGCTGAG 1560  
QY 1561 CTGTTTACAGAGAGCCCATGAGAGGCTGTGCTATCTCAGACCAAAAGCTCTCTTT 1620  
DB 1561 CTGTTTACAGAGAGCCCATGAGAGGCTGTGCTATCTCAGACCAAAAGCTCTCTTT 1620  
QY 1621 GCGGGCTCCGCGCTCAGCTGTGAGTGTGAGTGTGCGGCGCACTGATTAAGTATGCTGC 1680  
DB 1621 GCGGGCTCCGCGCTCAGCTGTGAGTGTGAGTGTGCGGCGCACTGATTAAGTATGCTGC 1680  
QY 1681 TGTGACAGACTGTCTCTGCGCGGAGCCGCTATTTGGCGTGAAGCTCAACACAGCGCG 1740  
DB 1681 TGTGACAGACTGTCTCTGCGCGGAGCCGCTATTTGGCGTGAAGCTCAACACAGCGCG 1740  
QY 1741 TGTGTGCGGCTGTGGCTGCTTGTGATCTTACTGATCCAGCATGTGATGACTCGGAC 1800  
DB 1741 TGTGTGCGGCTGTGGCTGCTTGTGATCTTACTGATCCAGCATGTGATGACTCGGAC 1800  
QY 1801 ACTTCAGGCAATTTCAACCTCCGAGAGTAAATACAGTGAAGGCCCACTCAAAAAAC 1860  
DB 1801 ACTTCAGGCAATTTCAACCTCCGAGAGTAAATACAGTGAAGGCCCACTCAAAAAAC 1860  
QY 1861 ATCAAGGTGTGTGGGCGACAGACCTGTGTGCTGCTGCGCACTCTCCCAACTTGACC 1920  
DB 1861 ATCAAGGTGTGTGGGCGACAGACCTGTGTGCTGCTGCGCACTCTCCCAACTTGACC 1920  
QY 1921 CTG 1923  
DB 1921 CTG 1923



Db 712 TAAATGAGGCCCCACACTCCATGAAAGAGATACCTGGCCCTTTGGCTCAAGAAC 771  
 QY 804 TCACTTTGTAGAGCTCTGCTATGATAGCTGAGAGTGTGGGAGCTTTCAGGGGGAGAG 863  
 Db 772 TCACTTTGTAGAGCTCTGCTATGATAGCTGAGAGTGTGGGAGCTTTCAGGGGGAGAG 831  
 QY 864 CAAGGTCTACTTCTCTTTCAGAGGAGCGGGAGTGGATCCGACTGCTATGCGGAGAGT 923  
 Db 832 CAAGGTCTACTTCTCTTTCAGAGGAGCGGGAGTGGATCCGACTGCTATGCGGAGAGT 891  
 QY 924 GGTGGCTCTGTGGCCCGCTGCTGCAAGGGGAGATATGGGGGGGGCCAGAGCCCTGCA 983  
 Db 892 GGTGGCTCTGTGGCCCGCTGCTGCAAGGGGAGATATGGGGGGGGCCAGAGCCCTGCA 951  
 QY 984 GAAATGAGACCACTGCTCTGAAAGGGCGGGCTGGACATGCTGCCCCGAAGTGGACCTTA 1043  
 Db 952 GAAATGAGACCACTGCTCTGAAAGGGCGGGCTGGACATGCTGCCCCGAAGTGGACCTTA 1011  
 QY 1044 CTTCACACCGAGCGAGCGATGACACACCTGCGAGACACCTCTGCGACAAACACCTCTT 1103  
 Db 1012 CTTCACACCGAGCGAGCGATGACACACCTGCGAGACACCTCTGCGACAAACACCTCTT 1071  
 QY 1104 CTTTGGGGTTTTCACACAGTGGGGGTGACATGATACCTGTCGGCCGCAATCTGTAGTACA 1163  
 Db 1072 CTTTGGGGTTTTCACACAGTGGGGGTGACATGATACCTGTCGGCCGCAATCTGTAGTACA 1131  
 QY 1164 GTTGAAGAGATCCAGCGGGGTGTTGAGGGCCCTATTAAGAGATACCATGAGAAAGCCCA 1223  
 Db 1132 GTTGAAGAGATCCAGCGGGGTGTTGAGGGCCCTATTAAGAGATACCATGAGAAAGCCCA 1191  
 QY 1224 GAAATGAGACCGCTACACTGACCTGTATACCCAGCCCTGCGGCTGCTGCTGATTAACAA 1283  
 Db 1192 GAAATGAGACCGCTACACTGACCTGTATACCCAGCCCTGCGGCTGCTGCTGATTAACAA 1251  
 QY 1284 CTGGCATCGGGCGGCGAGGCTACACAGCTCCCTGGAGCTACCCGAAACAATCCATCACTT 1343  
 Db 1252 CTGGCATCGGGCGGCGAGGCTACACAGCTCCCTGGAGCTACCCGAAACAATCCATCACTT 1311  
 QY 1344 CGTCAAGAGACCGCGCTGATGAGAGAGAGTGGGGCTGCTGCTGAGAGCCGCCCTGCT 1403  
 Db 1312 CGTCAAGAGACCGCGCTGATGAGAGAGAGTGGGGCTGCTGCTGAGAGCCGCCCTGCT 1371  
 QY 1404 CGTCAAGAGAGGCGACCACTTACCCACTGCTGCTGCGGAGCCGGGTTACAGAGCTTGATG 1463  
 Db 1372 CGTCAAGAGAGGCGACCACTTACCCACTGCTGCTGCGGAGCCGGGTTACAGAGCTTGATG 1431  
 QY 1464 AGCCACTATACAGTGTGCTCATTTGGACAGAGAGAGGCGGCTGCTCAAGGCTGTAG 1523  
 Db 1432 AGCCACTATACAGTGTGCTCATTTGGACAGAGAGAGGCGGCTGCTCAAGGCTGTAG 1491  
 QY 1524 CTTGGGGCCCTGGGCTACCTGATTTGAGAGAGCTGCTGTTTACAGAGAGCCCATGAG 1583  
 Db 1492 CTTGGGGCCCTGGGCTACCTGATTTGAGAGAGCTGCTGTTTACAGAGAGCCCATGAG 1551  
 QY 1584 AAGCTGTGCTATCTCAGAGCAAGAAAGCTGCTTTTGGCGGGCTCCGCTCTCAGCTGCT 1643  
 Db 1552 AAGCTGTGCTATCTCAGAGCAAGAAAGCTGCTTTTGGCGGGCTCCGCTCTCAGCTGCT 1611  
 QY 1644 GCAGCTGCGCGGTGGCGGCTGATTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1703  
 Db 1612 GCAGCTGCGCGGTGGCGGCTGATTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1671  
 QY 1704 GGACCCCTATTTGGCGGCTGAGAGCTCAACACAGCCGCTGTTGGCCGCTGCTGCTGCT 1763  
 Db 1672 GGACCCCTATTTGGCGGCTGAGAGCTCAACACAGCCGCTGTTGGCCGCTGCTGCTGCT 1731  
 QY 1764 TGGATCTTTACTGATCCAGCATGTGATGACTGGACACTTTCAGGCAATTTCCAACTCCG 1823  
 Db 1732 TGGATCTTTACTGATCCAGCATGTGATGACTGGACACTTTCAGGCAATTTCCAACTCCG 1791  
 QY 1824 TGGAGTAAAGTACAGTCAAGGCGCACTTCCAAAAAACAATCAAGGCTGGGGGGGACAGCA 1883  
 Db 1792 TGGAGTAAAG--AAAGTCAAGGCGCACT--CCCAAAAAACAATCAAGGCTGGGGGGGACAGCA 1848

QY 1884 CTTGCTGCTGCGCTGCGACCTCTCTCCAACTTGGCCCTGCGGCTCAACCCCGAG 1942  
 Db 1849 CTTGCTGCTGCGCTGCGACCTCTCTCCAACTTGGCCCGGAGCTGAGACCTTTGGGG 1907  
 RESULT 5  
 AAD28947  
 ID AAD28947 standard; cDNA; 3868 BP.  
 XX  
 AC AAD28947;  
 DT 07-MAY-2002 (first entry)  
 XX  
 DE Human MOL5a cDNA.  
 KW Secreted molecule; MOL5a protein; MOLX; cardiomyopathy; atherosclerosis;  
 KW diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;  
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;  
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;  
 KW immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;  
 KW HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;  
 KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;  
 KW ocular disease; growth disorder; depression; epilepsy; contraceptive;  
 KW vulnery; osteopathic; haemostatic; tranquiliser; antidepressant;  
 KW analgesic; vasotropic; hypotensive; gene therapy; chromosome 2; ss.  
 OS Homo sapiens.  
 XX  
 FH Key  
 FH Location/Qualifiers  
 FT 1..452  
 FT 5'UTR  
 FT 453..2954  
 FT CDS  
 FT /\*tag= a  
 FT /\*tag= b  
 FT /product= "Human MOL5a protein"  
 FT /transl\_except= (pos:1191..1193, aa:Asp)  
 FT 2955..3868  
 FT /\*tag= c  
 XX  
 PN WO200206339-A2.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 03-JUL-2001; 2001WO-US21249.  
 XX  
 PR 03-JUL-2000; 2000US-215854P.  
 PR 03-JUL-2000; 2000US-215856P.  
 PR 03-JUL-2000; 2000US-215902P.  
 PR 07-JUL-2000; 2000US-216585P.  
 PR 07-JUL-2000; 2000US-216586P.  
 PR 07-JUL-2000; 2000US-216722P.  
 PR 17-JUL-2000; 2000US-218922P.  
 PR 17-JUL-2000; 2000US-218922P.  
 PR 27-JUL-2000; 2000US-221285P.  
 PR 14-FEB-2001; 2001US-268734P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Spaderma SK, Tcherev V, Liu X, Shenoy S, Spytek K, Zethusen B;  
 PI Paturajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;  
 PI Alsobrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;  
 PI Padigaru M;  
 DR WPI: 2002-155038/20.  
 DR P-PSDB: AAE18213.  
 DR  
 XX  
 PT Nucleic acids encoding secreted polypeptides, designated MOLX  
 PT polypeptides, useful for treating a MOLX-associated disorder, e.g.  
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -  
 XX  
 PS Claim 3; Page 59-60; 223pp; English.

CC The patent discloses nucleic acid sequences encoding novel secreted  
CC molecule (MOL) polypeptides, designated MOLX polypeptides (i.e. a MOL  
CC protein where X is an integer from 1 to 8). Sequences of the invention  
CC are useful for treating or preventing a MOLX-associated disorder in  
CC humans. They are useful for treating or preventing cardiomyopathy,  
CC atherosclerosis and disorders related to cell signal processing and  
CC metabolic pathway modulation. The MOLX antibodies are useful for  
CC treating or preventing diabetes and disorders related to cell signal  
CC processing and metabolic pathway modulation. MOLX sequences are useful  
CC for the treatment or diagnosis of other MOLX-associated disorders, e.g.  
CC chromosomal disorders, albinism, anemia, liver cirrhosis, psoriasis,  
CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral  
CC thrombosis, ischemia, hypertension, systemic lupus erythematosus, immune  
CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,  
CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,  
CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,  
CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,  
CC ocular diseases, muscular diseases, growth disorders, loss of libido,  
CC stress, depression, pain and epilepsy. They are useful for preventing  
CC chemotherapy side effects and as contraceptives. Sequences of the  
CC invention are also useful for gene therapy. The present sequence  
CC is a cDNA encoding human semaphorin 4C-like protein, MOL5a. MOL5a  
CC gene is localised on chromosome 2.  
XX

SO Sequence 3868 BP; 717 A; 1106 C; 1207 G; 838 T; 0 other;

Query Match 76.1%; Score 1738; DB 24; Length 3868;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1782; Conservative 0; Mismatches 35; Indels 3; Gaps 2;

QY 123 CAGCAGGAGAGCTGGCGCTGAAGCTCAGACCGGGGGCTGGCGCTGCCCCACACCTGGGC 182  
DB |||||||  
QY 410 CAGCAGGAGAGCTGGCGCTGAAGCTCAGACCGGGGGCTGGCGCTGCCCCACACCTGGGC 469  
DB |||||||  
QY 183 TGTCTGGCTGTGGAGCAAGGCTGTGGGGCTGGGGCTATGGGGCTGAGTGTGGAA 242  
DB |||||||  
QY 470 TGTCTGGCTGTGGAGCAAGGCTGTGGGGCTGGGGCTATGGGGCTGAGTGTGGAA 529  
DB |||||||  
QY 243 CCTTGTGCCCTTAAGACAGTGTCTTGGGAGCTGGCCAGCTAGTACGGCGGTTCTC 302  
DB |||||||  
QY 530 CCTTGTGCCCTTAAGACAGTGTCTTGGGAGCTGGCCAGCTAGTACGGCGGTTCTC 589  
DB |||||||  
QY 303 CCAGACCGGCTACGAGAGCTTCTGACACTGACGCTAGCGAGCCACAGCGGCTTCTGTA 362  
DB |||||||  
QY 590 CCAGACCGGCTACGAGAGCTTCTGACACTGACGCTAGCGAGCCACAGCGGCTTCTGTA 649  
DB |||||||  
QY 363 CGTGGGCGCCGAGAGGCGCTTGTGCTTACAGTACGAGGCGCTGAGAGCTCAAGAGAC 422  
DB |||||||  
QY 650 CGTGGGCGCCGAGAGGCGCTTGTGCTTACAGTACGAGGCGCTGAGAGCTCAAGAGAC 709  
DB |||||||  
QY 423 GATCTCTGGAGAGGCGCCCGCTGGAGAGAAAGTATGTCAGAAAGGAGAAAGAA 482  
DB |||||||  
QY 710 GATCTCTGGAGAGGCGCCCGCTGGAGAGAAAGTATGTCAGAAAGGAGAAAGAA 769  
DB |||||||  
QY 483 CCAGACCGGCTACGAGAGCTTCTGACACTGACGCTAGCGAGCCACAGCGGCTTCTGTA 542  
DB |||||||  
QY 770 CCAGACCGGCTACGAGAGCTTCTGACACTGACGCTAGCGAGCCACAGCGGCTTCTGTA 829  
DB |||||||  
QY 543 CGTGTGGGACCTACGCGCTTCCAGCCCAAGTACAGTACGTAACATGCTCACTTAC 602  
DB |||||||  
QY 830 CGTGTGGGACCTACGCGCTTCCAGCCCAAGTACAGTACGTAACATGCTCACTTAC 889  
DB |||||||  
QY 603 TTTGAGAGCTGAGAGGTTTGAAGATGGAGAGGCAAGTGTCTTATGACCCAGCTAAGG 662  
DB |||||||  
QY 890 TTTGAGAGCTGAGAGGTTTGAAGATGGAGAGGCAAGTGTCTTATGACCCAGCTAAGG 949  
DB |||||||  
QY 663 CCATCTGGGCTTCTTGTGATGATGAGCTGACTCGGCGACACATCAACATCTTCTGG 722  
DB |||||||  
QY 950 CCATCTGGGCTTCTTGTGATGATGAGCTGACTCGGCGACACATCAACATCTTCTGG 1009  
DB |||||||  
QY 723 CACGAGACCATTAATCTGCGTACATGGGGCCCGACACATCAATCAATCAATCAAT 782  
DB |||||||  
QY 1010 CACGAGACCATTAATCTGCGTACATGGGGCCCGACACATCAATCAATCAATCAAT 1069  
DB |||||||

QY 783 GGCCTTTGGCTCAACGAACCTCCTAGCTTGTAGCTCTGCTATGTAAGTGTGG 842  
DB |||||||  
QY 1070 GGCCTTTGGCTCAACGAACCTCCTAGCTTGTAGCTCTGCTATGTAAGTGTGG 1129  
DB |||||||  
QY 843 CAGCTTACAGGGGGAGAGACAGACAGTCTTCTTCTTACAGGAGGGGAGTGTGAGTC 902  
DB |||||||  
QY 1130 CAGCTTACAGGGGGAGAGACAGTCTTCTTCTTCTTACAGGAGGGGAGTGTGAGTC 1189  
DB |||||||  
QY 903 CGACTCTATGCGGACAGAGT 962  
DB |||||||  
QY 1190 CGCTCTCTATGCGGACAGAGT 1249  
DB |||||||  
QY 963 GGGCGACGAGACCTTGCAGAGAAAGTGAACAGTTCCTGAAGGCGGCTGGCATGTC 1022  
DB |||||||  
QY 1250 GGGCGACGAGACCTTGCAGAGAAAGTGAACAGTTCCTGAAGGCGGCTGGCATGTC 1309  
DB |||||||  
QY 1023 TGCCTCCGAAGTGGAGAGCTTACTTCAACCAAGCTGAGGAGGATGACACCTGAGAGAC 1082  
DB |||||||  
QY 1310 TGCCTCCGAAGTGGAGAGCTTACTTCAACCAAGCTGAGGAGGATGACACCTGAGAGAC 1369  
DB |||||||  
QY 1083 CTCTGTGGCAACACACACCTTCTTGTGGGTTTTCAGACAGTGGGATGACATGTACCT 1142  
DB |||||||  
QY 1370 CTCTGTGGCAACACACACCTTCTTGTGGGTTTTCAGACAGTGGGATGACATGTACCT 1429  
DB |||||||  
QY 1143 GTGCGCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202  
DB |||||||  
QY 1430 GTGCGCATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1489  
DB |||||||  
QY 1203 GGAATACATGAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1262  
DB |||||||  
QY 1490 GGAATACATGAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1549  
DB |||||||  
QY 1263 GCTGTGCTGTGATTAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1322  
DB |||||||  
QY 1550 GCTGTGCTGTGATTAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1609  
DB |||||||  
QY 1323 ACCCGAACAATCTCTCAACTTCTGTCAGAGACACCGCTGATGAGAGAGAGTGGGGC 1382  
DB |||||||  
QY 1610 ACCCGAACAATCTCTCAACTTCTGTCAGAGACACCGCTGATGAGAGAGAGTGGGGC 1669  
DB |||||||  
QY 1383 TGGGTGAGAGGCGCGCGCTGCTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1442  
DB |||||||  
QY 1670 TGGGTGAGAGGCGCGCGCTGCTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1729  
DB |||||||  
QY 1443 CCGGGTTACAGAGCTTGTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1502  
DB |||||||  
QY 1730 CCGGGTTACAGAGCTTGTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1789  
DB |||||||  
QY 1503 CTGGCTGCTCAAGGCTGTGAGGCTGGGGCGCTGGGGCTGACCTGATTTGAGAGAGCTGAGCT 1562  
DB |||||||  
QY 1790 ATGGCTGCTCAAGGCTGTGAGGCTGGGGCGCTGGGGCTGACCTGATTTGAGAGAGCTGAGCT 1849  
DB |||||||  
QY 1563 GTTTCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1622  
DB |||||||  
QY 1850 GTTTCACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909  
DB |||||||  
QY 1623 CGGCTCCGCTCTCAAGCTGTGAGGCTGGCGCTGGCGGAGCTGATTAAGTATGCTCTG 1682  
DB |||||||  
QY 1910 CGGCTCCGCTCTCAAGCTGTGAGGCTGGCGCTGGCGGAGCTGATTAAGTATGCTCTG 1969  
DB |||||||  
QY 1683 TGCAGAGTGTGCTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1742  
DB |||||||  
QY 1970 TGCAGAGTGTGCTGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2029  
DB |||||||  
QY 1743 TGTGGCGGTGGTGGGCACTTTGATCTTATGATGATGATGATGATGATGATGATGATGAT 1802  
DB |||||||  
QY 2030 TGTGGCGGTGGTGGGCACTTTGATCTTATGATGATGATGATGATGATGATGATGATGAT 2089  
DB |||||||  
QY 1803 TTCAGGACATTTCAACCTCTGCGAGTAAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1862  
DB |||||||  
QY 2090 TTCAGGACATTTCAACCTCTGCGAGTAAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2146  
DB |||||||

QY 1863 CACGGTGGTGGCGGACACACCTGTCGTCCTGACCTCTCTCCAACTTGCCCT 1922  
 DB 2147 CAGGTGTGTGGCGGACACACCTGTCGTCCTGACCTCTCTCCAACTTGCCCA 2206  
 OY 1923 GCCGCACTCCACGCCGAGG 1942  
 DB 2207 TGGCCGCTGGACCTTTGGGG 2226

RESULT 6  
 AAD28949  
 ID AAD28949 standard; cDNA; 3112 BP.  
 AC AAD28949;  
 XX  
 DT 07-MAY-2002 (first entry)  
 XX  
 DE Human MO15c cDNA.

Secreted molecule; MO15c protein; MO1X; cardiomyopathy; atherosclerosis;  
 diabetes; chromosomal disorder; albinism; anaemia; psoriasis; scarring;  
 liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;  
 cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;  
 immune disease; ischaemia; immunodeficiency; rheumatoid arthritis; ulcer;  
 HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;  
 haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;  
 Parkinson's disease; Huntington's disease; muscular disease; stress;  
 ocular disease; growth disorder; depression; epilepsy; contraceptive;  
 vulnerability; osteopathic; haemostatic; tranquilliser; antidepressant;  
 analgesic; vasotropic; hypotensive; gene therapy; chromosome 2; 5s.

OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..103  
 FT 5'UTR  
 FT 104..2605  
 FT CDS  
 FT /\*tag- b  
 FT /\*product- "Human MO15c protein"  
 FT sig\_peptide  
 FT /\*tag- c  
 FT 164..2602  
 FT mat\_peptide  
 FT /\*tag- d  
 FT /\*product- "Human MO15c mature protein"  
 FT 2606..3112  
 FT 3'UTR  
 FT /\*tag- e

XX MO200206339-A2.  
 XX  
 XX 24-JAN-2002.  
 XX  
 XX 03-JUL-2001; 2001WO-US21249.  
 XX  
 XX 03-JUL-2000; 2000US-21585AP.  
 XX 03-JUL-2000; 2000US-21585BP.  
 XX 03-JUL-2000; 2000US-215902P.  
 XX 07-JUL-2000; 2000US-216585P.  
 XX 07-JUL-2000; 2000US-216586P.  
 XX 07-JUL-2000; 2000US-216723P.  
 XX 17-JUL-2000; 2000US-218622P.  
 XX 17-JUL-2000; 2000US-218992P.  
 XX 27-JUL-2000; 2000US-221285P.  
 XX 14-FEB-2001; 2001US-268734P.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Spaderne SK, Tchernov V, Liu X, Shenoy S, Spytek K, Zerhusen B;  
 XX Patunajan M, Taupier RJ, Rastelli L, Grosse WM, Szekeres ES;  
 XX Alsobrook J, Lapeley DM, Shen L, Burgess CE, Shinkets RA;  
 XX Padigaru M;  
 XX WPI; 2002-155038/20.  
 XX P-PSDB; MAE18215.

XX Nucleic acids encoding secreted polypeptides, designated MO1X  
 PT polypeptides, useful for treating a MO1X-associated disorder, e.g.  
 PT cardiomyopathy, atherosclerosis, diabetes and metabolic disorders -  
 PS Claim 3; Page 63-64; 223pp; English.  
 XX  
 XX The patent discloses nucleic acid sequences encoding novel secreted  
 CC molecule (MO1X) polypeptides, designated MO1X polypeptides (i.e. a MO1X  
 CC protein where X is an integer from 1 to 8). Sequences of the invention  
 CC are useful for treating or preventing a MO1X-associated disorder in  
 CC humans. They are useful for treating or preventing cardiomyopathy,  
 CC atherosclerosis and disorders related to cell signal processing and  
 CC metabolic pathway modulation. The MO1X antibodies are useful for  
 CC treating or preventing diabetes and disorders related to cell signal  
 CC processing and metabolic pathway modulation. MO1X sequences are useful  
 CC for the treatment or diagnosis of other MO1X-associated disorders, e.g.  
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,  
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral  
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune  
 CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,  
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,  
 CC haematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,  
 CC ulcer, Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC stress, depression, pain and epilepsy. They are useful for preventing  
 CC chemotherapy side effects and as contraceptives. Sequences of the  
 CC invention are also useful for gene therapy. The present sequence  
 CC is a cDNA encoding human semaphorin 4C-like protein, MO15c. MO15c  
 CC gene is localised on chromosome 2.  
 XX  
 XX Sequence 3112 BP; 559 A; 948 C; 957 G; 648 T; 0 other;  
 S0

Query Match 75.4%; Score 1722; DB 24; Length 3112;  
 Best Local Similarity 97.4%; Pred. No. 0;  
 Matches 1772; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

OY 123 CAGCAGGACAGCTGGCTGAAGCTCAGACCGGGGCGTCCGCAATGGCCCACTGGGC 182  
 DB 61 CAGCAGGACAGCTGGCTGAAGCTCAGACCGGGGCGTCCGCAATGGCCCACTGGGC 120  
 OY 183 TGTCTGGCTGCTGCAGCAGAGGCTGTGGGCGCTTGGGCTGAGTGTGTGGA 242  
 DB 121 TGTCTGGCTGCTGCAGCAGAGGCTGTGGGCGCTTGGGCTGAGTGTGTGGA 180  
 OY 243 CCTGTGGCGCGTAAGACAGTGTCTGTGGGAGCTGGCCAGGTAAGAGCGGTCTC 302  
 DB 181 CCTGTGGCGCGTAAGACAGTGTCTGTGGGAGCTGGCCAGGTAAGAGCGGTCTC 240  
 OY 303 CCAGACGGGCAATCCAGACTTCTGACACTGACCTGACGAGCCCACTGGCTTGTGA 362  
 DB 241 CCAGACGGGCAATCCAGACTTCTGACACTGACCTGACGAGCCCACTGGCTTGTGA 300  
 OY 363 CGTGGGCGCCGAGAGGCGCTGTTGCCCTCAGCATGGAGGCGCTGACGTCAAGAGC 422  
 DB 301 CGTGGGCGCCGAGAGGCGCTGTTGCCCTCAGCATGGAGGCGCTGACGTCAAGAGC 360  
 OY 423 GATCTCGTGGAGGCGCCCGTGGAGGAACAGTGAAGTTCAGAAAGGAAACAA 482  
 DB 361 GATCTCGTGGAGGCGCCCGTGGAGGAACAGTGAAGTTCAGAAAGGAAAGAA 420  
 OY 483 CCAGACGAGTGTCAATCACTTCAGACCTTCTGACAGGCTTCAATAGGCTCCCACTGA 542  
 DB 421 CCAGACGAGTGTCAATCACTTCAGACCTTCTGACAGGCTTCAATAGGCTCCCACTGA 480  
 OY 543 CGTCTGTGGCACTTACGCTTCCAGACCAAGTGCACCTGCAATGCTACCTTAC 602  
 DB 481 CGTCTGTGGCACTTACGCTTCCAGACCAAGTGCACCTGCAATGCTACCTTAC 540  
 OY 603 TTTGGAGCATGAGAGTGTGAAGTGAAGGGAAGTGTCCATGAGCCCAAGTAAAGG 662  
 DB 541 TTTGGAGCATGAGAGTGTGAAGTGAAGGGAAGTGTCCATGAGCCCAAGTAAAGG 600

QY 663 CCATCTGGCCCTCTTGTGGATGTAGCTGACTCGGCCACATCAACACTCTCTGGG 722  
 DB 601 CCATCTGGCCCTCTTGTGGATGTAGCTGACTCGGCCACATCAACACTCTCTGGG 660  
 QY 723 CACGGACCATTTATCTGGCGTACATGGGGGCCACCATCTCATGAAGACAGTACTT 782  
 DB 661 CACGGACCATTTATCTGGCGTACATGGGGGCCACCATCTCATGAAGACAGTACTT 720  
 QY 783 GGCCTTTTGGCTCAACGAACCTCACTTTGTAGCTCTGGCTATGTACCTGAGAGTGTGG 842  
 DB 721 GGCCTTTTGGCTCAACGAACCTCACTTTGTAGCTCTGGCTATGTACCTGAGAGTGTGG 780  
 QY 843 CAGCTTCAGGGGGGACAGACAGACAGAGTCTACTTCTTTCAGAGAGCGGGCAGTGAATC 902  
 DB 781 CAGCTTCAGGGGGGACAGACAGAGTCTACTTCTTTCAGAGAGCGGGCAGTGAATC 840  
 QY 903 CCACTGCTATGCGGACAGAGTGTGGCTCTGTGGCCGCTGCTCAAGGGGATATGGG 962  
 DB 841 CCACTGCTATGCGGACAGAGTGTGGCTCTGTGGCCGCTGCTCAAGGGGATATGGG 900  
 QY 963 GGGCGCAGGACCCCTGACAGAGAGTGCACAGTTCCTTCAAGGGCGGCTGGCATGTCTC 1022  
 DB 901 GGGCGCAGGACCCCTGACAGAGAGTGCACAGTTCCTTCAAGGGCGGCTGGCATGTCTC 960  
 QY 1023 TGCCCCGAAGTGGCAGCTCTACTTCAACCAAGCTGCAGGGCATGCACACCTGCAGAGAC 1082  
 DB 961 TGCCCCGAAGTGGCAGCTCTACTTCAACCAAGCTGCAGGGCATGCACACCTGCAGAGAC 1020  
 QY 1083 CTCTCTGGGACACACACCTCTTTGGGGTTTTCAGAGCAGTGGGGTGATGATGACTT 1142  
 DB 1021 CTCTCTGGGACACACACCTCTTTGGGGTTTTCAGAGCAGTGGGGTGATGATGACTT 1080  
 QY 1143 GTGGCCATCTGTGAGTACCACTTGGAGAGATTCACGGGGTGTTCAGAGGCCCTATTA 1202  
 DB 1081 GTGGCCATCTGTGAGTACCACTTGGAGAGATTCACGGGGTGTTCAGAGGCCCTATTA 1140  
 QY 1203 GGAGTACCATGAGAGAGCCAGAGTGGGACCGCTACACTGACCTGTACCCAGCCCTCG 1262  
 DB 1141 GGAGTACCATGAGAGAGCCAGAGTGGGACCGCTACACTGACCTGTACCCAGCCCTCG 1200  
 QY 1263 GCCTGGCTCTGATTAACAACTGATCGGCCACGGCTACACAGCTCCCTGGAGCT 1322  
 DB 1201 GCCTGGCTCTGATTAACAACTGATCGGCCACGGCTACACAGCTCCCTGGAGCT 1260  
 QY 1323 ACCCAGCAACATCTCTCACTTGTCAAGAGCAACCGGCTGATGAGAGACAGTGGGCC 1382  
 DB 1261 ACCCAGCAACATCTCTCACTTGTCAAGAGCAACCGGCTGATGAGAGACAGTGGGCC 1320  
 QY 1383 TGGGTGAGCGCGCCCTGCTGTGAGAGAGGGACCAACTTCAACCCAGCTGGGCCGA 1442  
 DB 1321 TGGGTGAGCGCGCCCTGCTGTGAGAGAGGGACCAACTTCAACCCAGCTGGGCCGA 1380  
 QY 1443 CCGGCTTACAGACTTGTATGAGCCACCTATACAGTGTCTTCAATTTGGACAGAGACGG 1502  
 DB 1381 CCGGCTTACAGACTTGTATGAGCCACCTATACAGTGTCTTCAATTTGGACAGAGACGG 1440  
 QY 1503 CTGGCTCTCAAGGCTGTGAGCCCTGGGGCCCTGGGTTCACTGATTTGAGAGACTCAAGT 1562  
 DB 1441 CTGGCTCTCAAGGCTGTGAGCCCTGGGGCCCTGGGTTCACTGATTTGAGAGACTCAAGT 1500  
 QY 1563 GTTTACCAAGAGAGCCATGAGAGAGCTGTGTATCTCAAGAGCAAGAGTCTCTTTC 1622  
 DB 1501 GTTTACCAAGAGAGCCATGAGAGAGCTGTGTATCTCAAGAGCAAGAGTCTCTTTC 1560  
 QY 1623 CGGCTCCGCTCTCAAGTGTGACCTGCGCCCTGGCCGACATGAGAGTATCGCTCTCG 1682  
 DB 1561 CGGCTCCGCTCTCAAGTGTGACCTGCGCCCTGGCCGACATGAGAGTATCGCTCTCG 1620  
 QY 1683 TGCAGACTGTGTCTTCGCCCCGGGACCCCTATTGCGCTGTGAGAGCTCAACACCGCGCTG 1742  
 DB 1621 TGCAGACTGTGTCTTCGCCCCGGGACCCCTATTGCGCTGTGAGAGCTCAACACCGCGCTG 1680  
 QY 1743 TGTGGCCGTGGGTGGCCACTTTGGATCTTTACTGATCCAGCATGTGATGACTCTGGACAC 1802

DB 1681 TGTGGCCGGGGTGGCCACTGTGATCTCTACTGATCCAGCATGTGATGATGACTCGGACAC 1740  
 QY 1803 TTTCAGGACATTTGCACACCTCCGGGGGACGTAAGTACAGTACAGGCCACCTGCCAAAAACAT 1862  
 DB 1741 TTTCAGGACATTTGCACACCTCCGGGGGACGTAAG--AAAGTAGGCCACT--CCCAAAAACAT 1797  
 QY 1863 CACGGTGTGGGGGACAGACACTGTGTGCTGCTGCCCTGCCACCTCTCCCAACTTGGCCCT 1922  
 DB 1798 CACGGTGTGGGGGACAGACACTGTGTGCTGCTGCCCTGCCACCTCTCCCAACTTGGCCCA 1857  
 QY 1923 GCCCGACTCAACCCCGAGG 1942  
 DB 1858 TGCCCGCTGGACCTTTGGG 1877  
 RESULT 7  
 AAD28948  
 ID AAD28948 standard; cDNA; 2558 BP.  
 XX  
 AC AAD28948;  
 XX  
 DT 07-MAY-2002 (first entry)  
 XX  
 DE Human MOL5b cDNA.  
 KW Secreted molecule; MOL5b protein; MOLX; cardiomyopathy; atherosclerosis;  
 KW diabetes; chromosomal disorder; albinism; anemia; psoriasis; scarring;  
 KW liver cirrhosis; neurodegeneration; osteoarthritis; organ rejection;  
 KW cerebral thrombosis; hypertension; systemic lupus erythematosus; asthma;  
 KW immune disease; ischemia; immunodeficiency; rheumatoid arthritis; ulcer;  
 KW HIV; human immunodeficiency virus; hepatitis; haemostatic disease; pain;  
 KW haematopoietic; thrombolytic; thrombocytopenia; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; muscular disease; stress;  
 KW ocular disease; growth disorder; depression; epilepsy; contraceptive;  
 KW vulvectomy; osteoporotic; haemostatic; tranquilizer; antidepressant;  
 KW analgesic; vasotrophic; hypotensive; gene therapy; chromosome 2; 9s.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..20  
 FT CDS /\*tag= a  
 FT /\*tag= b  
 FT /\*tag= c  
 FT 3'UTR 2523..2558  
 FT /\*tag= c  
 XX  
 PN MO200206339-A2.  
 XX  
 PD 24-JAN-2002.  
 PF 03-JUL-2001; 2001WO-US21249.  
 XX  
 PR 03-JUL-2000; 2000US-215854P.  
 PR 03-JUL-2000; 2000US-215856P.  
 PR 03-JUL-2000; 2000US-215902P.  
 PR 07-JUL-2000; 2000US-216585P.  
 PR 07-JUL-2000; 2000US-216586P.  
 PR 07-JUL-2000; 2000US-216722P.  
 PR 17-JUL-2000; 2000US-218622P.  
 PR 17-JUL-2000; 2000US-218992P.  
 PR 27-JUL-2000; 2000US-221285P.  
 PR 14-FEB-2001; 2001US-268734P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Spaderne SK, Tchiernev V, Liu X, Shenoy S, Spytek K, Zerhusen B;  
 PI Paturajan M, Taupier RJ, Rastelli L, Grose WM, Szekeres ES;  
 PI Alsbrook J, Lepley DM, Shen L, Burgess CE, Shinkets RA;  
 PI Padigaru M;  
 XX

DR WPI: 2002-15038/20.  
 DR P-PSDB: AAE18214.  
 XX Nucleic acids encoding secreted polypeptides, designated MOIX  
 PT polypeptides, useful for treating a MOIX-associated disorder, e.g.  
 XX cardiovascular, atherosclerosis, diabetes and metabolic disorders -  
 PS claim 3: Page 61; 223pp; English.  
 XX The patent discloses nucleic acid sequences encoding novel secreted  
 CC molecule (MOL) polypeptides, designated MOIX polypeptides (i.e. a MOL  
 CC protein where X is an integer from 1 to 8). Sequences of the invention  
 CC are useful for treating or preventing a MOIX-associated disorder in  
 CC humans. They are useful for treating or preventing cardiovascular,  
 CC atherosclerosis and disorders related to cell signal processing and  
 CC metabolic pathway modulation. The MOIX antibodies are useful for  
 CC treating or preventing diabetes and disorders related to cell signal  
 CC processing and metabolic pathway modulation. MOIX sequences are useful  
 CC for the treatment or diagnosis of other MOIX-associated disorders, e.g.  
 CC chromosomal disorders, albinism, anaemia, liver cirrhosis, psoriasis,  
 CC scarring, neurodegeneration, osteoarthritis, organ rejection, cerebral  
 CC thrombosis, ischaemia, hypertension, systemic lupus erythematosus, immune  
 CC diseases, immunodeficiency, HIV (human immunodeficiency virus), viral,  
 CC bacterial and fungal infections, hepatitis, rheumatoid arthritis, asthma,  
 CC hematopoietic, thrombolytic, haemostatic diseases, thrombocytopenia,  
 CC ulcers, Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC ocular diseases, muscular diseases, growth disorders, loss of libido,  
 CC stress, depression, pain and epilepsy. They are useful for preventing  
 CC chemotherapy side effects and as contraceptives. Sequences of the  
 CC invention are also useful for gene therapy. The present sequence  
 CC is a cDNA encoding human semaphorin 4C-like protein, MOIX5b. MOIX5b  
 CC gene is localised on chromosome 2.  
 XX  
 XX Sequence 2558 BP; 468 A; 788 C; 801 G; 501 T; 0 other;  
 SQ  
 Query Match 74.48; Score 1699; DB 24; Length 2558;  
 Best Local Similarity 97.38; Pred. No. 0;  
 Matches 1749; Conservative 0; Mismatches 45; Indels 3; Gaps 2;

DB	481	ATGGGAAGGCAAGTGTCCCTATGACCCAGCTAAGGGCCATGCTGACCTCTTGTGTGATG	540
QY	686	GTGAGCTGTACTGCGCCACACTCAACAACCTTCTGGGACGGAACCCATTATTCCTGGCTA	745
DB	541	GTGAGCTGTACTGCGCCACACTCAACAACCTTCTGGGACGGAACCCATTATTCCTGGCTA	600
QY	746	ACATGGGCCCCCACCACCTCATGAAAGACAGATGACCTGGCTTTTGGCTCAAGAACCTC	805
DB	601	ACATGGGCCCCCACCACCTCATGAAAGACAGATGACCTGGCTTTTGGCTCAAGAACCTC	660
QY	806	ACTTTGAGGCTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	865
DB	661	ACTTTGAGGCTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
QY	866	AGGCTACTTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	925
DB	721	AGGCTACTTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	780
QY	926	TGCTGCTGTGGCCGCTGTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	985
DB	781	TGCTGCTGTGGCCGCTGTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	840
QY	986	AGTGACACAGTTCCTGAAGGCGGCGTGGATGCTCTGCCCGCACTGGAGCTGACT	1045
DB	841	AGTGACACAGTTCCTGAAGGCGGCGTGGATGCTCTGCCCGCACTGGAGCTGACT	900
QY	1046	TCAACACAGTTCGACGCGATGACACACCTGACAGACCTCTGACACACACACCTTCT	1105
DB	901	TCAACACAGTTCGACGCGATGACACACCTGACAGACCTCTGACACACACACCTTCT	960
QY	1106	TTGGGGTTTTTCAAGACAGTGGGGTACATGATGATGATGATGATGATGATGATGATGATG	1165
DB	961	TTGGGGTTTTTCAAGACAGTGGGGTACATGATGATGATGATGATGATGATGATGATGATG	1020
QY	1166	TGGAAGAGATCCAGCGGCTGTTTGAAGGCGCCCTTAAGAGATACATGAGAGAACCCAGA	1225
DB	1021	TGGAAGAGATCCAGCGGCTGTTTGAAGGCGCCCTTAAGAGATACATGAGAGAACCCAGA	1080
QY	1226	AGTGGACCGCTACACTGACCTTACCCAGCCCTGCGGCTGCTGCTGCTTAACT	1285
DB	1081	AGTGGACCGCTACACTGACCTTACCCAGCCCTGCGGCTGCTGCTGCTTAACT	1140
QY	1286	GGCATGCGGCGCACGCGCTACACACGCTCCCTGAGCTACCCGACACATCTCAACTCG	1345
DB	1141	GGCATGCGGCGCACGCGCTACACACGCTCCCTGAGCTACCCGACACATCTCAACTCG	1200
QY	1346	TCAAGAAACACCCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1405
DB	1201	TCAAGAAACACCCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1260
QY	1406	TGAAGAAAGGACCAACCTTACCCAGCTGAGGCGGACCGGCTTACAGAGATTGATGAG	1465
DB	1261	TGAAGAAAGGACCAACCTTACCCAGCTGAGGCGGACCGGCTTACAGAGATTGATGAG	1320
QY	1466	CCACCTTATACAGTCTGCTTATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1525
DB	1321	CCACCTTATACAGTCTGCTTATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1380
QY	1526	TGGGGCCCTGGGCTTACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1585
DB	1381	TGGGGCCCTGGGCTTACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1440
QY	1586	GCCGTGTGTATCTCAAG	1645
DB	1441	GCCGTGTGTATCTCAAG	1500
QY	1646	AGCTGCCGTGGCGGACTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATG	1705
DB	1501	AGCTGCCGTGGCGGACTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATG	1560
QY	1706	ACCCCTATGCGCTGAGAGGCTCAACACACGCGCTGTGTGGCCGTGGGCGGCACTCTG	1765
DB	1561	ACCCCTATGCGCTGAGAGGCTCAACACACGCGCTGTGTGGCCGTGGGCGGCACTCTG	1620



QY 1766 GATCTTTACGTATCCAGCATGTGATGACCTGGGACACTTCAGGCATTTGCAACTCCGTG 1823

Db 1621 GATCTTACTATCCAGCATGTGATGACCTGGGACACTTCAGGCATTCGCAACTCCGTG 1680

OY 1826 GCAGTAAAGATACAGTCAGGCCCCACTNCCCAAAAAATCAGCGTGTGGCGGGGACAGACC 1885

Db 1681 GCAGTAAAG - AAACGACAGCCCCACT - CCCAAAAATCAGCGTGTGGCGGGGACAGACC 1737

OY 1886 TGGTGTCGCCCTGGCACCTCTCTCCAACTTGGGCGCTGGCGGACATCCAAACCCGAGG 1942

Db 1738 TGGTGTCGCCCTGGCACCTCTCTCCAACTTGGGCGGATGCCCGGTGGAGACTTTGGGG 1794

XX	RESULT 8
XX	ABN83983
ID	ABN83983 standard; DNA; 3776 BP.
XX	
AC	ABN83983;
XX	
DT	06-SEP-2002 (first entry)
XX	
DE	Human gene sequence #30.
XX	
KW	Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
CDS	232..2892
FT	/*tag= a
XX	
PN	WO200252005-A1.
XX	
PD	04-JUL-2002.
XX	
PE	20-DEC-2001; 2001WO-JP11217.
XX	
PR	22-DEC-2000; 2000JP-0389742.
XX	
PA	(KAZU-) KAZUSA DNA RES INST FOUND.
XX	(CELE-) CELESTAR LEXICO-SCI LTD.
PI	Obara O, Nagase T, Nakajima D;
DR	WPI: 2002-500762/53.
P-PSDB:	ABB97963.
PT	Genes and their expression products cloned from human cDNA libraries
PT	for treatment and diagnosis of diseases associated with their
PT	expression -
XX	
TS	Claim 1(a); Page 119-125; 238pp; Japanese.
XX	
CC	The invention relates to DNA encoding polypeptides directly cloned from
CC	cDNA libraries originating in adult whole brain, human tonsil, human
CC	adult hippocampus and human foetal whole brain. Polypeptides and
CC	polynucleotides of the invention may be used in the investigation of
CC	differential expression of the DNA sequences in normal subjects and
CC	disease patients. They may also be used in the production of antibodies,
CC	oligonucleotide probes and DNA chips for diagnosis and identification
CC	of drugs for treatment of diseases with which the DNA sequences are
CC	associated. The sequences given in records ABN83954-ABN83984 represent
CC	human gene sequences of the invention.
XX	
Sequence	3776 BP; 696 A; 1088 C; 1191 G; 801 T; 0 other;

Query Match	70.8%	Score 1617.6	DB 24	Length 3776
Best Local Similarity	98.7%	Pred. No. 0		
Matches 1652; Conservative	0	Mismatches 19	Indels 3	Gaps 2

QY 269 CTGGGAGCTGGCCACGCGTACGTACGGCGGCTTCTCCACAGCCGGCATCCAGACATTCTCTGA 328  
|||||

Db	498	CTGCAGAGCGTGGCCACGGAGTACGCGGGTTCTCCAGACCGGACATCCAGAGACTTCCTCA	553
QY	339	CACGTACGCGTGCAGGAGGCCCACTGGGGCTTCTGTACGTGGGGCCCGAGAGGCCCTGTGTTG	388
Db	554	CACGTACGCGTGCAGGAGGCCCACTGGGGCTTCTGTACGTGGGGCCCGAGAGGCCCTGTGTTG	613
QY	389	CCTTACGATGAGAGGCCCTGGAGCTGCAAGAGCCATCTCTGGAGAGGCCCCCTGGAGAG	448
Db	614	CCTTACGATGAGAGGCCCTGGAGCTGCAAGAGCCATCTCTGGAGAGGCCCCCTGGAGAG	673
QY	449	AGAACTAGTAGTATCCAGAAAGGGAAGAACACAGACCCAGAGGCTTCAACTCATCC	508
Db	674	AGAACTAGTAGTATCCAGAAAGGGAAGAACACACAGACCCAGAGGCTTCAACTCATCC	733
QY	509	GCTTCCTGCAGCCCTACAATGCGCTCCACCGCTGTACGTGTGGACCTACGCGCTTCAGC	568
Db	734	GCTTCCTGCAGCCCTACAATGCGCTCCACCGCTGTACGTGTGGACCTACGCGCTTCAGC	793
QY	569	CCAATGCACTTACGTCAACATGCTCACCTTCTACTTTGGAGCATGAGAGTTTGAAGATG	628
Db	794	CCAATGCACTTACGTCAACATGCTCACCTTCTACTTTGGAGCATGAGAGTTTGAAGATG	853
QY	629	GGAAGGGCAAGTGGCTCATGACCCAGGCTAAGGGCATGCTGSGGCTCTGTGGATGGATG	688
Db	854	GGAAGGGCAAGTGGCTCATGACCCAGGCTAAGGGCATGCTGSGGCTCTGTGTGGATGGT	913
QY	689	AGCTGTACTCGGGCCACACTCAACAACCTCTGGGCAAGGAAACCCATTATCCGCTTAACA	748
Db	914	AGCTGTACTCGGGCCACACTCAACAACCTCTGGGCAAGGAAACCCATTATCCGCGGTAA	973
QY	749	TGGGGCCCCACCACTTCATGAAAGACAGAGTACCTGGCCTTTTGGCTCAACGAACTCACT	808
Db	974	TGGGGCCCCACCACTTCATGAAAGACAGAGTACCTGGCCTTTTGGCTCAACGAACTCACT	1033
QY	809	TTGTAGGCTCTCCATATGACTGAGAGTGGGGAGCTCAAGGGGGACGACGACAAAG	868
Db	1034	TTGTAGGCTCTCCATATGACTGAGAGTGGGGAGCTTCAAGGGGGACGACGACAAAG	1093
QY	869	TCTACTTCTTCTTCAGGGAGCGGGCAGTGGAGTCCGACTGCTATCCGACAGCAGTGGTGG	928
Db	1094	TCTACTTCTTCTTTCAGGGAGCGGGCAGTGGAGTCCGACTGCTATCCGAGAGGAGTGGTGG	1153
QY	929	CTCGTGTGGCCGCTGTCTGCAAGGGCGCATATGGGGGGCGACAGAACCTCGACAGAGAGT	988
Db	1154	CTCGTGTGGCCGCTGTCTGCAAGGGCGCATATGGGGGGCGACAGAACCTCGCAGAGAGAGT	1213
QY	989	GGACACAGTCCCTGAAGCGCGGGCTGGGATGCTTGCCCGGAATCGGAGAGCTCACTCA	1048
Db	1214	GGACACAGTCTCTGAAGCGCGGGCTGGGATGCTTGCCCGGAATCGGAGAGCTCTCACTCA	1273
QY	1049	ACCAGCTGCAGGCGATGACACACCTCGAGGACCTCTGTGGACAAACACACCTTCTTGG	1108
Db	1274	ACCAGCTGCAGGCGATGACACACCTCGAGGACACCTCTGTGGACAAACACACCTTCTTGG	1333
QY	1109	GGGTTTTCAAGCACAGTGGGGTGACATGTACCTTCCGCCATCTGTGAAGTACCAAGTTGG	1166
Db	1334	GGGTTTTCAAGCACAGTGGGGTGACATGTACCTTCCGCCATCTGTGAAGTACCAAGTTGG	1393
QY	1169	AAGAGATCCAGGGGGTGTGTAAGGGGGCCCTAATAAGAGTACCATAGGAAGAACCCAGAGT	1222
Db	1394	AAGAGATCCAGGGGGTGTGTAAGGGGGCCCTAATAAGAGTACCATAGGAAGAACCCAGAGT	1453
QY	1229	GGAGACGGGTACACTAGCCTGTGACCACGCCCTGGGCTGGGTGCGGATTAATAACATGCG	1288
Db	1454	GGAGACGGGTACACTAGCCTGTGACCACGCCCTGGGCTGGGTGCGGATTAATAACATGCG	1513
QY	1289	ATCGGCGGCACGGGTACACACAGCTTCCGTGAAGCTTACCAGACAACTTCAACTCTCTCA	1348
Db	1514	ATCGGCGGCACGGGTACACACAGCTTCCGTGAAGCTTACCAGACAACTTCAACTCTCTCA	1573
QY	1349	AGAAACACCCGCTGATGAGAGGAGCAGTGGGGCTCTGGTGGAGACCGCCCTGCTGTGA	1408
Db	1574	AGAAACACCCGCTGATGAGAGGAGCAGTGGGGCTCTGGTGGAGACCGCCCTGCTGTGA	1633



QY 1409 AGAAGGCGACCACTTCACCCCTGGTGGCCGACCGGGTTACAGACCTTGATGAGACCA 1468  
|||||  
DB 1634 AGAAGGCGACCACTTCACCCCTGGTGGCCGACCGGGTTACAGACCTTGATGAGACCA 1693  
1469 CCTATACAGTGTGCTTATTTGGCAGAGAGAGGCGCTGGCTGCTCAAGGCTGTGAGCCCTG 1528  
1694 CCTATACAGTGTGCTTATTTGGCAGAGAGAGGCGCTGGCTGCTCAAGGCTGTGAGCCCTG 1753  
QY 1529 GGGCCGTGGGTTCACTGATTTAGAGAGCTGACAGCTGTTTGAACAGAGACCCATGAGAAC 1588  
|||||  
DB 1754 GGGCCGTGGGTTCACTGATTTAGAGAGCTGACAGCTGTTTGAACAGAGACCCATGAGAAC 1813  
QY 1589 TGGTGTCTATCTCAGAGCAAGAGCTGCTTTTGGCCGCTCCCGCTCTCAGCTGTGCAAC 1648  
|||||  
DB 1814 TGGTGTCTATCTCAGAGCAAGAGCTGCTTTTGGCCGCTCCCGCTCTCAGCTGTGCAAC 1873  
QY 1649 TGGCCCGTGGCGAGCTGATTAAGATATGCTGCTGTGAGAGCTGTGCTGCGCCGCGGAC 1708  
1874 TGGCCCGTGGCGAGCTGATTAAGATATGCTGCTGTGAGAGCTGTGCTGCGCCGCGGAC 1933  
QY 1709 CCTATTGCGCGCTGGAGCGCTGACACACAGCGCTGTGCGCGTGGTGGCCACTTTGGAT 1768  
|||||  
DB 1934 CCTATTGCGCGCTGGAGCGCTGACACACAGCGCTGTGCGCGTGGTGGCCACTTTGGAT 1993  
QY 1769 CTTTACTGATTCAGCATGTGATGACCTGCGACACTTCAGGCATTGGCAACCTCGTGACA 1828  
|||||  
DB 1994 CTTTACTGATTCAGCATGTGATGACCTGCGACACTTCAGGCATTGGCAACCTCGTGACA 2053  
QY 1829 GTAAAGTACAGTGCAGGCGCCACTGCCAATAATCAGAGGTGGTGGCGGCAACACTGG 1888  
|||||  
DB 2054 GTAAAGTACAGTGCAGGCGCCACTGCCAATAATCAGAGGTGGTGGCGGCAACACTGG 2110  
QY 1889 TGGTGGCGCTGCGACCTCTCTCCCAACTTTGGCCCTGCGGACTGCCAACCCGAGG 1942  
2111 TGGTGGCGCTGCGACCTCTCTCCCAACTTTGGCCCTGCGGACTGCCAACCCGAGG 2164

RESULT 9  
ABX92031  
ID ABX92031 standard; cDNA: 8095 BP.  
XX  
AC ABX92031:  
XX  
DT 08-MAY-2003 (first entry)  
XX  
DE Lung specific nucleic acid (LSNA) #73.  
XX  
KW Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;  
KW cancer monitoring; cancer staging; cancer imaging; lung cancer;  
KW non-cancerous diseases of the lung; transgenic animal; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200268633-A2.  
XX  
PD 06-SEP-2002.  
XX  
PE 21-NOV-2001; 2001WO-US43612.  
XX  
PR 22-NOV-2000; 2000US-252500P.  
XX  
PA (DIAD-) DIADEXUS INC.  
PI Macina RA, Recipon H, Chen S, Sun Y, Liu C:  
XX  
DR WPI: 2002-713376/77.  
XX  
PT New isolated human nucleic acid molecule and polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating lung  
PT cancer and non-cancerous diseases of the lung  
PS Claim 1; Page 221-226; 389pp; English.

XX The invention describes an isolated human nucleic acid (I) encoding any  
CC of 120-10-1533 residue amino acid sequences (S1), given in the  
CC specification, comprising any of 164 179-12421 base pair sequences (S2),  
CC given in the specification. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating lung cancer and non-cancerous diseases of the lung.  
CC They are also used for identifying lung tissue, monitoring and  
CC identifying and/or designing antagonists of the polypeptide of the  
CC invention, gene therapy, production of transgenic animals and production  
CC of engineered lung tissue for treatment and research. This sequence  
CC encodes a lung specific nucleic acid.  
XX  
SQ Sequence 8095 BP; 1899 A; 2246 C; 2050 G; 1900 T; 0 other;  
Query Match 61.08; Score 1394; DB 24; Length 8095;  
Best Local Similarity 97.68; Pred. No. 0;  
Matches 1448; Conservative 0; Mismatches 30; Indels 6; Gaps 3;  
QY 462 TATCCAGAAAGGGAACAACAACAGACCGAGTGTCTCACTTCATCCGCTTCGACGCC 521  
|||||  
DB 5887 TAAACCGACACTTCTTGCCCGCCAGACCGAGTGTCTCACTTCATCCGCTTCGACGCC 5946  
QY 522 CTACAAATGCTCCCACTGTACGTCTGTGGACCTACGCCCTTCCAGCCCAAGTGCACCTA 581  
|||||  
DB 5947 CTACAAATGCTCCCACTGTACGTCTGTGGACCTACGCCCTTCCAGCCCAAGTGCACCTA 6006  
QY 582 CGTCAATGCTCCCACTGTACGTCTGTGGACCTACGCCCTTCCAGCCCAAGTGCACCTA 641  
|||||  
DB 6007 CGTCAATGCTCCCACTGTACGTCTGTGGACCTACGCCCTTCCAGCCCAAGTGCACCTA 6066  
QY 642 TCCCTATGACCCAGCTTAAGGGCCATGCTGCGCTTCTGTGATGATGAGTGTGAGTGGC 701  
6067 TCCCTATGACCCAGCTTAAGGGCCATGCTGCGCTTCTGTGATGATGAGTGTGAGTGGC 6126  
DB 702 CACACTCAACACTTCTCGGGCAGGAAACCATTTATCTGTGCTTAACATGGGGCCCCACCA 761  
|||||  
DB 6127 CACACTCAACACTTCTCGGGCAGGAAACCATTTATCTGTGCTTAACATGGGGCCCCACCA 6186  
QY 762 CTCCATGAAAGACAGTACTGCTGCGCTTCTGTGCTCAACGAACTCTGTAAGCTCTGC 821  
6187 CTCCATGAAAGACAGTACTGCTGCGCTTCTGTGCTCAACGAACTCTGTAAGCTCTGC 6246  
DB 822 CTATGTACTGAGAGTGTGGGCACTTACAGGGGAGCAGACAGCAAGTCTACTTCTT 881  
6247 CTATGTACTGAGAGTGTGGGCACTTACAGGGGAGCAGACAGCAAGTCTACTTCTT 6306  
QY 882 CAGGAGCGGGCAGTGGAGTCCGACTGTATGCGCGAGCAGTGGTGGCTCTGTGGCCCG 941  
6307 CAGGAGCGGGCAGTGGAGTCCGACTGTATGCGCGAGCAGTGGTGGCTCTGTGGCCCG 6366  
QY 942 TGTCTGCAAGGGGCGATATGGGGGCGCAGGAGCCCTGCAAGAGTGGACACAGTCTT 1001  
6367 TGTCTGCAAGGGGCGATATGGGGGCGCAGGAGCCCTGCAAGAGTGGACACAGTCTT 6426  
DB 1002 GAAGGCGCGCTGCGATGCTTGCCTCGAATGCGAGCTCTACTTCAACACAGCTGACGC 1061  
6427 GAAGGCGCGCTGCGATGCTTGCCTCGAATGCGAGCTCTACTTCAACACAGCTGACGC 6486  
QY 1062 GATGCAACCTCTGAGAGACACCTGCTGCAACACACACTTCTTTGGGGTTTTTCAAGC 1121  
6487 GATGCAACCTCTGAGAGACACCTGCTGCAACACACACTTCTTTGGGGTTTTTCAAGC 6546  
DB 1122 ACAGTGGGCTGACATGTACTGTGCGGCGATGTGAGTACAGAGTGGAAAGATCCACGC 1181  
6547 ACAGTGGGCTGACATGTACTGTGCGGCGATGTGAGTACAGAGTGGAAAGATCCACGC 6606  
QY 1182 GGTGTTTGGAGGCCCTTATAAGAGTACATAGAGAACCCAGAAAGTGGAGCCCTACAC 1241  
6607 GGTGTTTGGAGGCCCTTATAAGAGTACATAGAGAACCCAGAAAGTGGAGCCCTACAC 6666  
DB 1242 TGACCCCTTACCCAGCCCTCGGCTGGCTGTGCTTAAACAATGGCATCGGCGCACGG 1301  
|||||

Db	6667	TGACCCTGTACCCACAGCCCTGGGCTGTGCTGTGATTTAAACAACCTGGATGGGCGCACGG	6726
QY	1302	CTACACCAAGCTCCCTGGAGCTACCCGACAAATCTCTAACTTGTCTCAAGAAGCACCCGCT	1361
Db	6727	CTTACACCAAGCTCCCTGGAGCTACCCGACAAATCTCTAACTTGTCTCAAGAAGCACCCGCT	6786
QY	1362	GATGGAGGAGCAGGTGGGGGCTCGGTGGAGCCGCGCCCTGCTCGTGAAGAAAGGACACAA	1421
Db	6787	GATGGAGGAGCAGGTGGGGGCTCGGTGGAGCCGCGCCCTGCTCGTGAAGAAAGGACACCA	6846
QY	1422	CTTACACCAAGCTCGTGGCCGACCGGGTTACAGACTTGTATGAGACCACTATACAGTGT	1481
Db	6847	CTTACACCAAGCTCGTGGCCGACCGGGTTACAGACTTGTATGAGACCACTATACAGTGT	6906
QY	1482	GTTCATTGGGACAGAGAGGAGCTGGCTGCTCAAGAGCTGTGAGCTTGAGGCGCTGGGGCT	1541
Db	6907	GTTCATTGGGACAGAGAGGAGCTGGCTGCTCAAGAGCTGTGAGCTTGAGGCGCTGGGGCT	6966
QY	1542	CCTGATTGAGGAGCTGAGCTGTTTGAACAAGAGCCCATGAGAAAGCTGTGCTATCTCA	1601
Db	6967	CCTGATTGAGGAGCTGAGCTGTTTGAACAAGAGCCCATGAGAAAGCTGTGCTATCTCA	7026
QY	1602	GAGCAAG---AAGTGCCTTTTGGCCGGCTCCCGCTACAGCTGGTGAAGCTGGCCGTGGC	1658
Db	7027	GAGCAAGGTAAGCTGCTCTTGGCCGGCTCCCGCTACAGCTGGTGAAGCTGGCCGTGGC	7086
QY	1659	CGACTGCATTAAGATGCTGCTGTGAGAGACTGTGCTCTGCGCCGCGGGAGCCCTATTTGCGC	1718
Db	7087	CGACTGCATGAAGATGCTGCTGTGAGAGACTGTGCTCTGCGCCGCGGGAGCCCTATTTGCGC	7146
QY	1719	CTGGAGGGTCAACACAGCCGCTGTGGCCGCTGGGGGCGACATTGGATCTTACATGAT	1778
Db	7147	CTGGAGGGTCAACACAGCCGCTGTGGCCGCTGGGGGCGACATTGGATCTTACATGAT	7206
QY	1779	CCAGCATGTGTGACTCTCGACACTTTCAGGCACTTTCGAACCTCCGTGGCAGTAAGATACA	1838
Db	7207	CCAGCATGTGTGACTCTCGACACTTTCAGGCACTTTCGAACCTCCGTGGCAGTAAG--AAA	7264
QY	1839	GTCAGGCCCACTNCCCAAAAACATCAGCGTGTGGCGGCGACAGACTGTGCTGCTCGCTG	1898
Db	7265	GTCAGGCCCACT--CCCAAAAACATCAGCGTGTGGCGGCGACAGACTGTGCTGCTCGCTG	7323
QY	1899	CCACCTCTCCCTCAAACTTGGCCCTGCGCCGACCTCCAAACCCGAGG	1942
Db	7324	CCACCTCTCCCTCAAACTTGGCCCTGCGCCGACCTCCAAACCTTGGGG	7367
RESULT 10			
ABT13397			
ID	ABT13397	standard; DNA; 8144 BP.	
AC	ABT13397;		
DT	30-JAN-2003	(first entry)	
DE	Breast specific related polynucleotide SEQ ID No 112.		
KW	Cytostatic; BSP-agonist; BSP-antagonist; vaccine; gene therapy; cancer;		
KW	metastatic; breast cancer; breast specific; human; ds.		
OS	Homo sapiens.		
XX	W0200277232-A2.		
XX	03-OCT-2002.		
XX	21-NOV-2001; 2001WO-US43815.		
XX	22-NOV-2000; 2000US-252509P.		
XX	(DIAD-) DIADEXUS INC.		
XX	Salceda S, MacIina RA, Reclpon H, Pluta J, Sun Y, Liu C;		

Query Match	Best Local Similarity	61.0%	Score 1394	DB 25	Length 8144	
Matches 1448	Conservative	0	Mismatches	30	Indels	6
						Gaps 3
XX	WP1: 2003-018927/01.					
DR						
XX	New isolated nucleic acid molecule, useful for treating breast cancer,					
PT	and diagnosing or monitoring the presence of metastases of breast					
PT	cancer in a patient					
XX						
PS	Claim 1, Page 256-260; 377pp; English.					
XX						
CC	The invention relates to a novel isolated nucleic acid molecule					
CC	comprising: a sequence encoding a sequence comprising 11-1518 amino					
CC	acids; a sequence comprising 190-8144 bp; or a sequence that selectively					
CC	hybridises to, or having at least 60% identity with the 11-1518 amino					
CC	acid or 190-8144 nt sequence. The polypeptide and the nucleic acid are					
CC	useful for treating breast cancer, and diagnosing or monitoring the					
CC	presence of metastases of breast cancer in a patient. The polynucleotides					
CC	of the invention can be used to treat disorders by gene therapy. This					
CC	polynucleotide represents a breast specific related sequence of the					
CC	invention.					
XX						
S0	Sequence 8144 BP; 1900 A; 2247 C; 2053 G; 1944 T; 0 other;					
Query Match	Best Local Similarity	61.0%	Score 1394	DB 25	Length 8144	
Db	Matches 1448	Conservative	0	Mismatches	30	Indels 6; Gaps 3;
OY	462 TATCCAGAAAGGAGAAACAACACAGACGAGATGCTTCAACTTCATCCGCTTCGACGCC	521				
Db	5936 TAAACGGGACACCTTCTGCGCCCGACAGCGAGTGTTCACACTTATCCGCTTCGACGCC	5995				
OY	522 CTACAAATGCTCCCACTGTACGTCTGTGGCACTACGCTTCCAGCCCAAGTGCACCTA	581				
Db	5996 CTACAAATGCTCCCACTGTACGTCTGTGGCACTACGCTTCCAGCCCAAGTGCACCTA	6055				
OY	582 CGTCAACATGCTCACTCTCACTTGGAGACATGAGAGATTGAATGGAGGAGCAAGTG	641				
Db	6056 CGTCAACATGCTCACTCTCACTTGGAGACATGAGAGATTGAATGGAGGAGCAAGTG	6115				
OY	642 TCCCTATGACCCAGCTTAAGGGCCATGCTGCGCTTCTTGTGGATGTGAGCTGTACTCGGC	701				
Db	6116 TCCCTATGACCCAGCTTAAGGGCCATGCTGCGCTTCTTGTGGATGTGAGCTGTACTCGGC	6175				
OY	702 CACACTCAACAACTTCCGGGGCAGGAAACCATATCTGTGTAACTATGGGGCCCCACCA	761				
Db	6176 CACACTCAACAACTTCCGGGGCAGGAAACCATATCTGTGTAACTATGGGGCCCCACCA	6235				
OY	762 CTCATGAGAGACAGATGACTGCGCTTGTGGCTCAAGAACCTCACTTGTAGGCTCTGC	821				
Db	6236 CTCATGAGAGACAGATGACTGCGCTTGTGGCTCAAGAACCTCACTTGTAGGCTCTGC	6295				
OY	822 CTATGTATCTGAGAGTGTGGCAGCTTACGCGGGGACGACGACAAAGTCTTACTTCTTCTT	881				
Db	6296 CTATGTATCTGAGAGTGTGGCAGCTTACGCGGGGACGACGACAAAGTCTTACTTCTTCTT	6355				
OY	882 CAGGAGAGGGCAGTGGAGTCCGACTGCTATGAGCAGAGAGTGGTGGCTGTGGGCGG	941				
Db	6356 CAGGAGAGGGCAGTGGAGTCCGACTGCTATGAGCAGAGAGTGGTGGCTGTGGGCGG	6415				
OY	942 TGTCTGCAGGGCAGTATGGGGGCGCAGGACCTTGACAGAGAAATGTGACCAACGTTCT	1001				
Db	6416 TGTCTGCAGGGCAGTATGGGGGCGCAGGACCTTGACAGAGAAATGTGACCAACGTTCT	6475				
OY	1002 GAAGGCGGGCTGGCATATGCTGCCCCGAATGTGGCAGTCTACTTCAACAGCGTGCAGGC	1061				
Db	6476 GAAGGCGGGCTGGCATATGCTGCCCCGAATGTGGCAGTCTACTTCAACAGCGTGCAGGC	6535				
OY	1062 GATGCACACCTCTGAGGACACCTCTCTG6ACACACACCTTCTTGGGGTTTTCAGC	1121				
Db	6536 GATGCACACCTCTGAGGACACCTCTCTG6ACACACACCTTCTTGGGGTTTTCAGC	6595				
OY	1122 ACAAGTGGGGTACATATGATCTGTGGGCTATTTGATGATACCAATTTGGAAGATTCACGC	1181				
Db	6596 ACAAGTGGGGTACATATGATCTGTGGGCTATTTGATGATACCAATTTGGAAGATTCACGC	6655				



OY	607	GAGCATGAGAGTTGGAAGATGGGAAGGGCAAGTGTCCCTATGACCACAGCTAAAGGACCAT	666
Db	541	GACCTGCAGAATTTGAGAGATGGGAAGGGTAATATCCATATAGACCACACTAAGGCTCAC	600
OY	667	GCTGGCCCTCTTGTGGATGTGTAGCTGTACTGGGCCACACTCAGACACTTCCCTGGGACAG	726
Db	601	ACCGGACTCCCTGTGGAGCGGTGAGCTGTACTAGCCACACACTCATTAAACAGAGTACTGGCT	660
OY	727	GAACCCATTATCCCTCCGTAACTGGGGGCCCCACACTCCATTAACAGAGAGTACCTGGGCC	786
Db	661	GAGCCGGTATTCCTTCGATACATGGGGACCCACCCACTCCATTAACAGAGAGTACTGGCT	720
OY	787	TTTTGGCTCAAGAACCTCACTTTGTAGGCTCTGCTTAATGTACTGAGAGTGTGGGACAC	846
Db	721	TTTTGGCTGAATGAACCCCACTTTGTAGGCTCTGCTTTGTCTTGAGAGTGTGGGAAGC	780
OY	847	TTCAAGGGGGAGAGACAGACAAAGSTCTACTTCTTTCAGAGGAGCGGGCAGTGGATGCCAG	906
Db	781	TTCAAGGGGAGAGAGACAAAGTCTACTTCTTTCAGAGGAGCGGGCAGTGGATGAATGAC	840
OY	907	TGCTATCCCGAGACAGTGTGGCTGTGTGGCCCTGTCTCGCAAGGGCGATATGGGGGAC	966
Db	841	TGCTATTCGAGACAGTGTGGCTGTGTGGCGAAGATCTGTAAAGGTGTACATTTGGGGGGA	900
OY	967	GCACGGACCTCTGCAGAGAGATGGACACAGTTCTGTAAAGGGCGGCTGGCATGTCTGCC	1022
Db	901	GCACGGACGCTGCAGAGAGAAATGGACGACGTTCCCTGAAAGGCTGGTGTGTGCTCAGCC	960
OY	1027	CCGAATCGGACGCTACTACTTCACCAAGCTGGCAGGGAGATGCACACCTGTGAGGACACCTCC	1088
Db	961	CCTGACTGTGAAGGTCTACTTCAACACAGCTGAAGGGGTGTGCACACCTCTGGGGGCGCTCT	1022
OY	1087	TGGCACAACACCACTCTTGTGGGTTTTCAGAGCAGTGGGAGTGTACATGTACTGTCTG	1144
Db	1021	TGGCACAACACCACTCTTGTGGGTTTTCAGAGGCGCATGGGGAGATATGTAGACTGTCT	1080
OY	1147	GCGATCTGTGATTCACCAATGTGAAGAGATCCAGCGGGGTGTGAAGGGCCCTAATAAGAG	1200
Db	1081	GCACTTGTGTGATTCACCAATGTGAAGAGATCCAGCAAGTGTGAAGGGTCCCTACAAAGAG	1144
OY	1207	TACCATAGAGAACCCAGAAAGTGGGACCGCTACACTGACCCCTGTACCCAGCCCTCGGCT	1266
Db	1141	TACAGTGAAGAACCCAGAAAGTGGGACCGCTATACTGACCCGCTACCCAGCCCTCGGCT	1200
OY	1267	GGCTGTGTGCATTTAACAACTGGCATGTGGGCCACAGGCTACACAGCTCCCTGGAGCTACCC	1322
Db	1201	GGTGTGTATATCAACAACTGGCATGGACGAACTGTGATACCAAGTTCCTGTGAACCTGTCCG	1266
OY	1327	GACCAACATCTCAACCTTGTCAAGAGAGACCCGCTGATGAGAGAGCAAGTGGGGGCGCTCGG	1388
Db	1261	GACCAACACCTCTCAACCTTATCAAGAGAGACCCCGTGTGATGAGAGACCAAGTGAAGCTCGG	1322
OY	1387	TGAGACCGCCCGCTGTGCTGTGAAGAGGGCACCACTTCAACCACCTGTGGTGGCGACGG	1444
Db	1321	TTGGGCCCCCGCTCACTGTGTGAAGAGAACTAACTTACACACAGTGGTGGCCCAACAGG	1388
OY	1447	GTTACAGAGACTTGAATGGAGCCACGATATCAATGTGCTCATTTGGCAGAGAGAGGGCGAG	1506
Db	1381	GTTCCAGAGGCTGTGATGTGTGCCACCTATACAGTGTGTTCATTTGGTACAGAGAGTGGCTGG	1444
OY	1507	CTGTCTCAAGGCTGTGAGGCTGTGGGCGCTGTGGGTTCACTGATTTGAGAGACTGCAGCTGTT	1566
Db	1441	CTGTCTGAAGGCTGTGAGGCTGTGGGCGCTGTGATTCACATGTGTGAGAGAACTGGAAGTGT	1500
OY	1567	GACCAAGAGCCCATGAGAAAGCTGTGCTATCTCAGAGCAAGAAAGTGTCTTTTGGCGGC	1622
Db	1501	GACCAAGAGCCCATGAGAAAGTCTGTGTGTCTCTCAGAGCAAGAAAGTGTCTTTTGGCTGGC	1566
OY	1627	TCCCGGCTCTCAGCTGTGATCGCTGGCGGCTGGCGGCGCATTAAGATATAGCTCCCTGTGCA	1688
Db	1561	TCCCGGCTCTCAGCTGTGATCGCTGTGGCCGACTGCACAAAGTACCGCTTTCTGTGTGA	1622

OY	1687	GACGTGTGCTCGCCCGGAGCCCTATGTGCCTGGAGCGTCAACACAGCGCGTGTG	1746
DB	1691	GACGTGTGCTCGCCGAGGAGCCCTTACTGTGCTCGAATGTCAACACAGCGCGTGTG	1680
OY	1747	G---CCGTGGTGCCCACTTTGGATCTTTACTGATCCAGCATGTGATGACTCGGACACT	1803
DB	1691	GCCACCAACCAATGGTGTGCGTTCGGGGTCCCTTTCTGTGTCCAACTGTGGGCAACTTGGACACT	1740
OY	1804	TCAGGCAATTTGCACATCCGTGGGAGTAAATACAGTCAGGCGCCACTGCCCAAAACATC	1853
DB	1741	TCAAAGATGTGTAAACAGTATGGCATTTAA---AAAGTCAGATCTATTTCCCAACACATC	1797
OY	1864	ACGGTGTGTGGCGGACACAGACCTGTGTGCTGCCCTGCACACTCTCTCTCCAACTGTGGCCGTG	1923
DB	1798	ACCGTGTGTGTGAGGACACAGACCTGTGTCTACCTGCACACTCTCTCTCAATTTGGCCCAT	1857
OY	1924	CCGCACTCCAAACCCCG	1939
DB	1858	GCCCACTGTGACCTTTCG	1873

RESULT 12
AAF29451
ID AAF29451 standard; cDNA; 2811 BP.

DT 03-APR-2001 (first entry)  
XX

DE Human TANGO 276 cDNA.

KW Human INTERCEPT 211; INTERCEPT 297; TANGO 276; TANGO 282; TANGO 325  
KW TANGO 331; TANGO 332; cytostatic; antiinflammatory; antiarrhythmic;  
KW antiproliferic; gene therapy; cancer; inflammatory disorder;  
KW cardiac disorder; arrhythmia; skin disorder; psoriasis; ss.  
KW

**Homo sapiens.**

PN WO200100638-A2

PD 04-JAN-2001

PF 16-JUN-2000; 2000WO-US16658.

PR 29-JUN-1999; 99US-0342364.

PA (MILL-) MILLENNIUM PHARM INC.

PI K1rst SJ, Holtzman DA, Fraser CC, Sharp JD, Barnes TS;

DR WPI; 2001-061966/07.

**XX**

PT prevention of cancers, inflammatory disorders,

XX

PS Claim 1; Page 307; 372pp; English.

CC The present sequence is given in a specification relating to isolated

CC TANGO 292, TANGO 325, TANGO 331 and TANGO 332. These proteins are useful

CC regulate cellular processes e.g. growth, proliferation, survival,

CC diagnosed, prevented and treated by administration of these polypeptides,

CC disorders, cardiac disorders e.g. arrhythmia, and skin disorders e.g.

CC express the proteins in a host cell in gene therapy applications.

CC of the proteins in target cells. Fragments of the nucleic acid molecules

CC polymerase chain reaction (PCR) primers.

SQ Sequence 2811 BP; 501 A; 812 C; 888 G; 610 T; 0 other;

Query Match 52.3% Score 1193.6; DB 22; Length 2811;

Best Local Similarity 74.6%; Pred. No. 8.3e-271;

Matches 1716; Conservative 0; Mismatches 423; Indels 161; Gaps 11;

QY 118 CGCAGCAGCAGACAGCTGGCTTGAAGCTCAGAGACCCGGGCGCTCCGACATGGCCCCACAC 177  
DB 10 CGGCTCGGGGAGAGCTGGCTTGAAGCTCAGAGACCCGGGCGCTCCGACATGGCCCCACAC 69  
QY 178 TGGGCTGTGTGCTGTCTGTGGAGCAAGGCTGTGGGCTGGGCTATGGGGCTGAGGTGG 237  
DB 70 TGGGCTGTGTGCTGTCTGTGGAGCAAGGCTGTGGGCTGGGCTATGGGGCTGAGGTGG 129  
QY 238 TGGAACTTGTGGCGGCTTGAAGAGTGTCTGTGGGAGCTGGGCGGCTGATACGGCGG 297  
DB 130 TGGAACTTGTGGCGGCTTGAAGAGTGTCTGTGGGAGCTGGGCGGCTGATACGGCGG 189  
QY 298 TTTCCAGACCGGCTATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGGCTT 357  
DB 190 TTTCCAGACCGGCTATCCAGAGACTTCTGACACTGACGCTGACGAGGCCACTGGGCTT 249  
QY 358 CTGTACCTGGGCGCCGAGAGGCGCTTGTGCTTTCAGACATGAGAGCCCTGGAGCTGCA 417  
DB 250 CTGTACCTGGGCGCCGAGAGGCGCTTGTGCTTTCAGACATGAGAGCCCTGGAGCTGCA 309  
QY 418 GGAGCGATCTGCTGGGAGGCGCGCTTGAAGAGAGTGTATCCAGAAAGGAGAG 477  
DB 310 GGAGCGATCTGCTGGGAGGCGCGCTTGAAGAGAGTGTATCCAGAAAGGAGAG 369  
QY 478 AACACACAGACCGAGTGTCTTCACTTATCCGTTTCTGACGAGCCCTACATGCTCCAC 537  
DB 370 AACACACAGACCGAGTGTCTTCACTTATCCGTTTCTGACGAGCCCTACATGCTCCAC 429  
QY 538 CTGTACCTGTGGACCTTACGCTTCCAGGCCAAGTGCACATACGTC----- 585  
DB 430 CTGTACCTGTGGACCTTACGCTTCCAGGCCAAGTGCACATACGTCGTCGTC 489  
QY 586 ----- 585  
DB 490 CTCTACCTGGGTGCCCGACGCCCGCTCCTCCTCCTCTCTGTGACTGTGGATGT 549  
QY 586 -----AAC 589  
DB 550 GGCCACAGAGCCCTGGCCCTTAAGCATCTCTCATACCTCTCTCTGTCTTGAACA 609  
QY 590 TGTCTACCTTCACTTTGGAGCATGAGAGTTGAAGATGGAGGGCAAGTGTCTTATG 649  
DB 610 TGTCTACCTTCACTTTGGAGCATGAGAGTTGAAGATGGAGGGCAAGTGTCTTATG 669  
QY 650 ACCCAGTAAGGGGCAATGCTGGGCTTGTGTGAGTGTGAGCTGTACTCGGCCACACTCA 709  
DB 670 ACCCAGTAAGGGGCAATGCTGGGCTTGTGTGAGTGTGAGCTGTACTCGGCCACACTCA 729  
QY 710 ACAACTTCTGTGGAGCAAGCAATATCTGCTTAACATGGGGCGCCACACTGCAATGA 769  
DB 730 ACAACTTCTGTGGAGCAAGCAATATCTGCTTAACATGGGGCGCCACACTGCAATGA 789  
QY 770 AGACAGAGTACCTGGCTTTTGGCTCAACGAACCTCACTTTGAGGCTCTGCTATGTAC 829  
DB 790 AGACAGAGTACCTGGCTTTTGGCTCAACGAACCTCACTTTGAGGCTCTGCTATGTAC 849  
QY 830 CTGAGAGTGTGGGAGCTTCAACGGGGAGAGAGAGAGTACTTCTTTCAGGGAGC 889  
DB 850 CTGAGAGTGTGGGAGCTTCAACGGGGAGAGAGAGAGTACTTCTTTCAGGGAGC 909  
QY 890 GGGAGTGTGAGTCCGACTGTATGCGGAGAGTGTGGTGTGCTGTGGCGGCTGTCTCA 949  
DB 910 GGGAGTGTGAGTCCGACTGTATGCGGAGAGTGTGGTGTGCTGTGGCGGCTGTCTCA 969  
QY 950 AGGGCATATGGGGGGGCGACGAGACCTGACAGAGAAAGTGAACACGTTCTGAGGGCC 1009

DB 970 AGGGCATATGGGGGGGCGACGAGACCTGACAGAGAAAGTGAACACGTTCTTGAAGGCGC 1029  
QY 1010 GGGGAGATGTCTGGCCCCGAACTGGAGAGCTTCACTTCAACACCTGTGAGGCGATGCA 1069  
DB 1030 GGGTGGAGTGTCTGGCCCCGAACTGGAGAGCTTCACTTCAACACCTGTGAGGCGATGCA 1089  
QY 1070 CCGTGGAGAGACCTTCTGGGCAACACACACCTTCTTGGGGTGTTCACACAGTGGG 1129  
DB 1090 CCGTGGAGAGACCTTCTGGGCAACACACACCTTCTTGGGGTGTTCACACAGTGGG 1149  
QY 1130 GTGACATGTACTGTGTGGCCATCTGTGTAGTACCAAGTTGGAAGAGATCCAGCGGTGTTG 1189  
DB 1150 GTGACATGTACTGTGTGGCCATCTGTGTAGTACCAAGTTGGAAGAGATCCAGCGGTGTTG 1209  
QY 1190 AGGGCCCCATTAAGAGAGTACCATGAGAGAGCCGAGAGTGGGAGCGGTACACCTGACCTG 1249  
DB 1210 AGGGCCCCATTAAGAGAGTACCATGAGAGAGCCGAGAGTGGGAGCGGTACACCTGACCTG 1269  
QY 1250 TACCCAGCCCTGCGGCTGTGCTGTGATTAACAACCTGGCATCGGCGCCAGGCTACACCA 1309  
DB 1270 TACCCAGCCCTGTGTTGTGTGATGCTGTGCCAGC -CCGCCATGCGGCGGCTTACACATGCT 1328  
QY 1310 GCTCCCTGGAGCTTACCCGACAAACATCTCAACTTGTTCAGAGAGCACCCGCTGATGAGG 1369  
DB 1329 TTTGAGAGAGACAGGGGGCGGCGCTGTGCTGAAGGCTTACCTTGTGGCTGTGGCGAG 1388  
QY 1370 AGCAGTGTGGGCGCTGCGTGTGAGAGCCGCCCTGCTGTGTAACAAAGGAGCACCACTTACCC 1429  
DB 1389 GCGCGTGTGTACCTTGGAGGCGCGGCGCCCTGTGAAACCTGGGGCTGTGTGGCTGG 1448  
QY 1430 ACCTGTGTGGCGGAGGCTTACAGAGACTTGTGTGAGAGCCACTTATACAGTGTGTTCATG 1489  
DB 1449 CGGTGTGTGGCTGTGGGGCTGTGTGTGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1504  
QY 1490 GCACAGAGAGAGGCTGTGCTGTCAAGGCTGTGAGGCTGTGGGCGCTGTGGTTCACCTGATG 1549  
DB 1505 CGCGGCGGCTGTGGGAGAGAGCTGTGAGAGAGAGGCGCAAGGCTGTGAGAGAGAGCT -TGG 1562  
QY 1550 AGGAGCTGACCTGTTTGAACAG 1609  
DB 1563 TGTACCCCTGTGAGAGTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1610 AGTGTCTTGTGGCGCTCCGCTCTCACTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1669  
DB 1621 ACCAGATGAGAACTTGTGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680  
QY 1670 AGTATGCTCTGTGTGAGAGAGT 1729  
DB 1681 AGTATGCTGTGGAGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740  
QY 1730 ACACAGGCGGCTGT 1789  
DB 1741 CCCAGGCGGCTGT 1800  
QY 1790 TGACCTGGAACCTTCAAGGCAATTTGCAACCTTCCGAGAGTAAAGTATGATGAGGCGCAC 1849  
DB 1801 TGCCAAATGTATGCTGT 1860  
QY 1850 TNCCTAAAGATCAAGT 1909  
DB 1861 CCTGCTGATGTGCGGATGAACTGAGAGCAACTGTGAGCAAGCCAGCA ----- 1912  
QY 1910 CCAACTGTGGCGCTGT 1969  
DB 1913 -----CTGCCGATCTCAACCCCGAGAGTATCATGATGAGAGGAAACCCCGAC 1961  
QY 1970 GCGGTGCGGAGAGAGCTGT 2029  
DB 1982 GCGGTGCGGAGAGAGCTGT 2021  
QY 2030 AGGAGCATGAGGAG 2089

Db	2022	AGGCACATGGACGACGACACCTGCTCTGTCTGAGACACATACCTGCCACACACCCACCCGGC	2081
Qy	2090	CATGAGACCTGCTCTGCTGCTAGACAGCGGACACTG-CACCTGGGTGG-TCACAGGCGACC	2147
Db	2082	CATGAGACACCTGCTCTGCTGCTGACGACGGGACACTGCCATCTGGTGTGGCTCACACAGGGCAC	2141
Qy	2148	AG-CTGCAGAGGACATCTTCCTCTCTCTGTAATCAGACAGACAGCGGGGACCCACACCG	2206
Db	2142	AGCTTCGACAGAGGACATCTTCCTCTCTCTCTCTGTGTAATCAGACAGACAGCGGGGACCCACCG	2201
Qy	2207	CCAAAA-TTTTCAAGCAGAAAG-TTMAAGATGTGTGTTGTTGATTTGACATGTGTTG	2264
Db	2202	CCAAAACTTTTCAAGCAGAAAGTTTCAAGATGTGTGTTGTTGATTTGACATGTGTTG	2261
Qy	2265	TGTGTGTGTGTATGTGTGTG 2284	
Db	2262	TGTGTGTGTGTATGTGTG 2281	
RESULT 13			
ABNS9626			
ID	ABNS9626	standard; cDNA: 2813 BP.	
AC	ABNS9626;		
XX	28-JUN-2002	(first entry)	
DE	Novel human coding sequence SEQ ID NO: 37.		
XX	Human; antihemetic; vulnerary; antiinflammatory; immunomodulator;		
KW	antifertility; cerebroprotective; cytosstatic; rheumatic; gene therapy;		
KW	neuroprotective; antiparkinsonian; protein therapy; EST;		
KW	expressed sequence tag; gene; ss.		
OS	Homo sapiens.		
XX	WO200222660-A2.		
PN	21-MAR-2002.		
PD	10-SEP-2001; 2001WO-US26015.		
PF	11-SEP-2000; 2000US-0659671.		
XX	(HXSE-) HXSEQ INC.		
PA	Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;		
XX	Xue AJ, Yang Y, Wehrman T, Drmanac RT;		
PI	WPI: 2002-292408/73.		
PI	P-PSDB: ABB97213.		
DR	An isolated polynucleotide for treating diseases associated with its		
PT	encoded polypeptide such as cancer and multiple sclerosis -		
PT	Claim 1; SEQ ID NO 37; 509pp; English.		
PS	The present invention provides the protein and coding sequences of 444		
XX	novel human proteins. These were isolated from expressed sequences tags		
CC	hemaatopoiests e.g. to treat aplastic anemia, to help tissue regrowth		
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat		
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat		
CC	infertility, to regulate haemostasis or thrombolysis e.g. to treat		
CC	stroke and cancer, to screen for drugs, to treat inflammatory conditions		
CC	e.g. Rheumatoid arthritis, and to treat nervous system disorders e.g.		
CC	Parkinson's disease. The present sequence is a coding sequence of the		
XX	invention.		
XX	Sequence 2813 BP; 524 A; 805 C; 885 G; 597 T; 2 other;		
Query Match	47.98; Score 1094.2; DB 24; Length 2813;		
Boot Local Similarity	98.08; Pred. No. 2e-247;		

	Matches	1128:	Conservative	0:	Mismatches	20:	Indels	3:	Gaps	2:
OY	792	GCTCAAGCAACCTCACTTTGTAGGCTCTGCCATATGACCTGAGAGTGTGGGACCTTAC	851							
Db	16	GCTCAAGCAACCTCACTTTGTAGGCTCTGCCATATGACCTGAGAGTGTGGGACCTTAC	75							
OY	852	GGGGGAGACACAAAGGTCACTCTTCTTCACAGGACCGGCACTGGACCTCCACTA	911							
Db	76	GGGGGAGACACAAAGGTCACTCTTCTTCACAGGACCGGCACTGGACCTCCACTA	135							
OY	912	TGCCAGAGAGGTGTGGCTCTGTGTGGCCCGTGTGCAAGGGCCATATGAGGGGACG	971							
Db	136	TGCCAGAGAGGTGTGGCTCTGTGTGGCCCGTGTGCAAGGGCCATATGAGGGGACG	195							
OY	972	GACCCCTGAGAGGAAGTGTGACACAGCTTCTGAAGGGCGGCTGGCATGCTTGCCCGAA	1031							
Db	196	GACCCCTGAGAGGAAGTGTGACACAGCTTCTGAAGGGCGGCTGGCATGCTTGCCCGAA	255							
OY	1032	CTGGCAGCTCACTTCAACCCAGCTGCAAGGCAATCACACCCTGAGAGCACTCTCTGCA	1091							
Db	256	CTGGCAGCTCACTTCAACCCAGCTGCAAGGCAATCACACCCTGAGAGCACTCTCTGCA	315							
OY	1092	CAACACCACTCTCTTTGGGGTTTTTCAAGCACAGTGGGGTGCATGTACCTGTGCGCAT	1151							
Db	316	CAACACCACTCTCTTTGGGGTTTTTCAAGCACAGTGGGGTGCATGTACCTGTGCGCAT	375							
OY	1152	CTGTGAGTACCAAGTTGGAGAGATCCACGCGGCTTTTGAAGGGCCCTATGAAGAGTACCA	1211							
Db	376	CTGTGAGTACCAAGTTGGAGAGATCCACGCGGCTTTTGAAGGGCCCTATGAAGAGTACCA	435							
OY	1212	TGAGGAAGCCCGAAGTGTGGACCGCTCACTGACCCCTGTACCCAGCCCTGGGCTGCTC	1271							
Db	436	TGAGGAAGCCCGAAGTGTGGACCGCTCACTGACCCCTGTACCCAGCCCTGGGCTGCTC	495							
OY	1272	GTGATTAACAACATGGGATGGGGCCCAAGGCTAACCAAGCTCCCTGGAGCTACCCGCA	1331							
Db	496	GTGATTAACAACATGGGATGGGGCCCAAGGCTAACCAAGCTCCCTGGAGCTACCCGCA	555							
OY	1332	CATCCTCAACTCGTCAAGAAGCACCCGCTGATGGAAGAGCAGTGGGGCTCGGTGGAG	1391							
Db	556	CATCCTCAACTCGTCAAGAAGCACCCGCTGATGGAAGAGCAGTGGGGCTCGGTGGAG	615							
OY	1392	CCGCCCTCTCTGTGAAGAAGGGCACCCACTTACCCACTCGTGTGGCCGACCGGGTTAC	1451							
Db	616	CCGCCCTCTCTGTGAAGAAGGGCACCCACTTACCCACTCGTGTGGCCGACCGGGTTAC	675							
OY	1452	AGACTTGATGGAGCCACCTATACAGTGTCTTTCATTTGGGACAGAGAGCGGCTGGTGT	1511							
Db	676	AGACTTGATGGAGCCACCTATACAGTGTCTTTCATTTGGGACAGAGAGCGGCTGGTGT	735							
OY	1512	CAAGGCTGTGAGCCTTGGGGCCCTGGGTTACACTGATTGAGAGCTGCAGCTGTTTGACA	1571							
Db	736	CAAGGCTGTGAGCCTTGGGGCCCTGGGTTACACTGATTGAGAGCTGCAGCTGTTTGACA	795							
OY	1572	GGAGCCCATGGAAGCCGTGGTGTATATCAGAGCAAGAGCTGCTCTTTGGCCGCTCCG	1631							
Db	796	GGAGCCCATGGAAGCCGTGGTGTATATCAGAGCAAGAGCTGCTCTTTGGCCGCTCCG	855							
OY	1632	CTCTCAGCTGTGACGTGCCCCGTGGCCGACATGATAAAGTATGCTCTGTGCACTG	1691							
Db	856	CTCTCAGCTGTGACGTGCCCCGTGGCCGACATGATAAAGTATGCTCTGTGCACTG	915							
OY	1692	TGTCCTGGCCCGGAGACCCCTATTGCGCTGGAGGCTCAACACAGCCGCTGTGGCCGT	1751							
Db	916	TGTCCTGGCCCGGAGACCCCTATTGCGCTGGAGGCTCAACACAGCCGCTGTGGCCGT	975							
OY	1752	GGGTGGCCACTTTGGATCTTACTAGATCCAGCATGTATGACCTCGGACACTTACGCAT	1811							
Db	976	GGGTGGCCACTTTGGATCTTACTAGATCCAGCATGTATGACCTCGGACACTTACGCAT	1035							
OY	1812	TTTGCAACCTCGTGTGGCAGTAAAGATACGTACAGCCCACTNCCCAAAAATCACGGTGT	1871							
Db	1036	CTTGCAACCTCGTGTGGCAGTAAAG -- AAGTCAAGGCCCACT - CCCAAAAATCACAGGTGT	1092							

QY 1872 GCGGGGACAGACCTGTCGTCCTCCACCTCTCTCCACTTGGCCCTGCCGACTC 1931  
DB 1093 GGGGGGACAGACCTGTCGTCCTCCACCTCTCTCCACTTGGCCCTGCCGACTC 1152  
QY 1932 CAACCCGAGG 1942  
DB 1153 GACCTTGGGG 1163

RESULT 14  
ABN85379  
ID ABN85379 standard; DNA: 1024 BP.

AC ABN85379;  
DT 21-OCT-2002 (first entry)

DE Partial human NOV2, semaphorin-like protein, coding sequence.

XX Human; NOV2; cytosolic; Cardiac; Anti-inflammatory; Immunosuppressive;  
KM Antiallergic; Hemostatic; Anti-HIV; Antidiabetic; Anorectic;  
KM Antiparasitic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;  
KM Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;  
KM Gene Therapy; NOV; cancer; heart disease; inflammation;  
KM autoimmune disorder; allergy; blood disorder; AIDS; diabetes;  
KM obesity; asthma; IgA nephropathy; cirrhosis; arthritis;  
KM Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;  
KM wasting disorder; semaphorin-like protein; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH 33..1022  
FT CDS

FT /tag- a  
FT /partial  
FT /product- "Partial NOV2"  
FT /note- "This sequence only encodes residues 1 to  
330 of NOV2, no stop codon given"

XX MO2025704-A2.

XX 18-JUL-2002.

XX 09-JAN-2002; 2002MO-US00554.

XX 09-JAN-2001; 2001US-260417P.

XX 10-JAN-2001; 2001US-260831P.

XX 28-FEB-2001; 2001US-27238P.

XX 09-MAR-2001; 2001US-274876P.

XX 18-APR-2001; 2001US-284704P.

XX (CURA-) CURAGEN CORP.

XX Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Splytek KA;  
PI Zhong M, Gargoli EA, Burgess CE, Patturajan M, Vernet CAM;  
PI Taylor S, Tcherny VT, Miller CE, Guo X, Boldog FL, Grosse RM;  
PI Alsbrook JP, Gerlach V, Edinger S, Rothenberg ME, Elleman K;  
PI Maddougall J, Malynker U, Millet I, Peyman J, Smithson G;  
PI Gunther E, Stone DJ;

XX WPI: 2002-590674/63.

XX P-PSDB: ABB98402.

XX NOVX polypeptides and encoding polynucleotides, useful for preventing  
or treating NOVX-associated disorders e.g. cancer, inflammation, or  
Alzheimer's disease, and in chromosome mapping, tissue typing or  
pharmacogenomics  
Claim 9; Page 20; 358pp; English.

XX The present sequence is a partial coding sequence for a NOV protein. The  
CC NOV proteins and coding sequences are useful for treating or preventing

CC NOV-associated disorders or in the manufacture of a medicament for  
CC treating the disorders, such as cancer, heart disease, inflammation,  
CC autoimmune disorders, allergies, blood disorders, AIDS, diabetes,  
CC obesity, asthma, IgA nephropathy, cirrhosis, arthritis, Alzheimer's  
CC disease, infections (e.g. bacterial, viral, parasitic), stroke, muscular  
CC dystrophy, epilepsy, and other wasting disorders associated with chronic  
CC diseases. NOV2 is a semaphorin like protein.

XX Sequence 1024 BP; 206 A; 296 C; 320 G; 202 T; 0 other;

Query Match 40.6%; Score 928.4; DB 24; Length 1024;  
Best Local Similarity 92.3%; Pred. No. 1.5e-208;  
Matches 1023; Conservative 0; Mismatches 1; Indels 84; Gaps 1;

QY 134 CTGGCTGAAGCTCAGAGCCGGGCGTGGCCGATGGCCCACTGGGCTGTGGCTG 193  
DB 1 CTGGCTGAAGCTCAGAGCCGGGCGTGGCCGATGGCCCACTGGGCTGTGGCTG 60  
QY 194 TGGCAGCAGAGCTGTGGGCGCTGGGCAATGGGCTGAGCTGTGGAACTGTGCCG 253  
DB 61 TGGCAGCAGAGCTGTGGGCGCTGGGCAATGGGCTGAGCTGTGGAACTGTGCCG 120  
QY 254 GTAAGACAGTCTCTTGTGGGAGCTGGCCACAGGTATGACGGCGTTCTCCAGACCGCA 313  
DB 121 GTAAGACAGTCTCTTGTGGGAGCTGGCCACAGGTATGACGGCGTTCTCCAGACCGCA 180  
QY 314 TCCAGACCTCTCTACACTGACGTGACGGAGCCCACTGGGCTCTGTAGCTGGGCGCC 373  
DB 181 TCCAGACCTCTCTACACTGACGTGACGGAGCCCACTGGGCTCTGTAGCTGGGCGCC 240  
QY 374 GAGAGGCCCTTTGCTTTCAGCATGAGGCGCTTGAGCTGCAAGAGCCATCTCTGGG 433  
DB 241 GAGAGGCCCTTTGCTTTCAGCATGAGGCGCTTGAGCTGCAAGAGCCATCTCTGGG 300  
QY 434 AGGCCCCCGTGGAGAGAGAGTATCCAGAAAGGAGAAACCAACCAACCGAGT 493  
DB 301 AGGCCCCCGTGGAGAGAGAGTATCCAGAAAGGAGAAACCAACCAACCGAGT 360  
QY 494 GCTTCACTTATCCGCTTCTGACAGCCCTTCAATGCTCCACCTGTAGCTGTGGCA 553  
DB 361 GCTTCACTTATCCGCTTCTGACAGCCCTTCAATGCTCCACCTGTAGCTGTGGCA 420  
QY 554 CCTAGCCCTTCCAGCCCAAGTGCACCTTCAATGCTCCACCTTCTTGGAGCATG 613  
DB 421 CCTAGCCCTTCCAGCCCAAGTGCACCTTCAATGCTCCACCTTCTTGGAGCATG 480  
QY 614 GAGAGTTGAAGATGGGAGGAGCAAGTGTCCCTATGACCCAGCTATAGGCGCATGCTGGCC 673  
DB 481 GAGAGTTGAAGATGGGAGGAGCAAGTGTCCCTATGACCCAGCTATAGGCGCATGCTGGCC 540  
QY 674 TTCTTGTGATGTGAGCTGTACTGCGGCACACCAACTTCTCTGGGAGGAAACCA 733  
DB 541 TTCTTGTGATGTGAGCTGTACTGCGGCACACCAACTTCTCTGGGAGGAAACCA 600  
QY 734 TTATCTGCTGTACATGGGGGCCCCACACTCCATGAAAGAGAGTACTGTGCTTTGGC 793  
DB 601 TTATCTGCTGTACATGGGGGCCCCACACTCCATGAAAGAGAGTACTGTGCTTTGGC 660  
QY 794 TCAACGAACCTCATTGTAGGCTGTGCTATGTAACCTGAGAGTGTGGGAGCTTACGG 853  
DB 661 TCAACG----- 666  
QY 854 GGGAGCAGCAGAGTCTACTTCTTCAAGGAGCGGAGTGTAGTCCAGTGTATG 913  
DB 667 -----GGAGCGGGGAGTGTAGTCCAGTGTATG 696  
QY 914 CCGAGCAGTGTGTGCTGTGTGCCCCGTGTCTGCAAGGCGATATGGGGGCGCCACGCA 973  
DB 697 CCGAGCAGTGTGTGCTGTGTGCCCCGTGTCTGCAAGGCGATATGGGGGCGCCACGCA 756  
QY 974 CCTGTGAGAGAGATGTGACCAAGCTTCTGAAAGGCGGCTGTGATGCTCTCCCGCAACT 1033  
DB 757 CCTGTGAGAGAGATGTGACCAAGCTTCTGAAAGGCGGCTGTGATGCTCTCCCGCAACT 816



QY 1034 GGCACCTCTACTTCAACGAGTGCAGGCGATGACACACCTTCGAGGACACCTCTGGCACA 1093  
 DB 817 GGCACCTCTACTTCAACGAGTGCAGGCGATGACACACCTTCGAGGACACCTCTGGCACA 876  
 QY 1094 ACACACCTCTCTGGGTTTTTCAAGACAGTGGGGTGACATGTAACCTGTGGGCATCT 1153  
 DB 877 ACACACCTCTCTGGGTTTTTCAAGACAGTGGGGTGACATGTAACCTGTGGGCATCT 936  
 QY 1154 GTGACTACAGTGGAGAGATCCAGCGGTTGTGAGAGGCGCCCTATAGAGATACATG 1213  
 DB 937 GTGACTACAGTGGAGAGATCCAGCGGTTGTGAGAGGCGCCCTATAGAGATACATG 996  
 QY 1214 AGGAAGCCAGAGTGGAGCGCTACAC 1241  
 DB 997 AGGAAGCCAGAGTGGAGCGCTACAC 1024  
 RESULT 15  
 AAF30193  
 ID AAF30193 standard; cDNA; 967 BP.  
 AC AAF30193;  
 XX  
 XX 30-APR-2001 (first entry)  
 DE Clone 14998905.0.65 encoding SECP6.  
 XX  
 XX SECP6; secreted protein; human; diagnosis; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 166..904  
 FT /\*tag= a  
 FT sig\_peptide 166..225  
 FT /\*tag= b  
 FT mat\_peptide 226..901  
 FT /\*tag= c  
 PN WO200105971-A2.  
 PD 25-JAN-2001.  
 XX  
 PF 20-JUL-2000; 2000MO-US19890.  
 PR 20-JUL-1999; 99US-0144722.  
 PR 29-NOV-1999; 99US-0167785.  
 PR 19-JUL-2000; 2000US-0619252.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinketsu RA, Fernandes E;  
 XX  
 DR MPI: 2001-091973/10.  
 DR P-PSDB: AAB20160.  
 XX  
 PT Now polypeptide designated SECP, its encoding nucleic acid and its  
 PT immunospecific antibody, useful for diagnosing, preventing and treating  
 PT SECP-associated disorders such as cancer -  
 XX  
 XX Claim 8; Fig 6; 124pp: English.  
 XX  
 CC The present sequence is that of a portion of clone 14998905.0.65  
 CC encoding novel human protein SECP6 (see AAB20160), which is  
 CC predicted to localise in the microbody (peroxisome). The clone was  
 CC obtained from lymphoid tissue. SECP6 shows identity to murine  
 CC semaphorin-4c, a Type I membrane protein widely expressed in the  
 CC nervous system during development. The invention provides 9 novel  
 CC SECP secreted proteins (see AAB20155-63), nucleic acids encoding  
 CC them (see AAF30188-96), antibodies, mutants or fragments. These  
 CC can be used to detect, treat or prevent an SECP-associated disorder,  
 CC to screen for predisposition to such a disorder, and to identify an

CC agent that modulates the expression or activity of SECP.

XX Sequence 967 BP; 191 A; 274 C; 288 G; 214 T; 0 other;

Query Match 36.8%; Score 841.4; DB 22; Length 967;

Best Local Similarity 99.9%; Pred. No. 4,4e-186;

Matches 842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCCCTTCTCAGACCTCCCTGCTGATGTGTAAGCGGGGTTGGGGTTCTGACGGCT 60  
 DB 1 CGGGCCCTTCTCAGACCTCCCTGCTGATGTGTAAGCGGGGTTGGGGTTCTGACGGCT 60  
 QY 61 ATTGTCTGCGCTGGGGAAAGGGAGACAGCGCGGAGCCGGAACCTCCGCTGACAGCGCGCC 120  
 DB 61 ATTGTCTGCGCTGGGGAAAGGGAGACAGCGCGGAGCCGGAACCTCCGCTGACAGCGCGCC 120  
 QY 121 ACCACGAGACAGTGGGCTGAGGCTGAGAGCGGGGGTGGGCGCATGGCCCAACATGG 180  
 DB 121 ACCACGAGACAGTGGGCTGAGGCTGAGAGCGGGGGTGGGCGCATGGCCCAACATGG 180  
 QY 181 GCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 181 GCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 QY 241 AACCTTTGCGCGCTGAGACAGTGTCTCTGCGGAGCTGCGCCAGCTGATGCGCGCTTC 300  
 DB 241 AACCTTTGCGCGCTGAGACAGTGTCTCTGCGGAGCTGCGCCAGCTGATGCGCGCTTC 300  
 QY 301 TCCGAGACCGGATCCAGAGACTTCTGACAGTGTCTCTGCGGAGCTGCGCCAGCTGATG 360  
 DB 301 TCCGAGACCGGATCCAGAGACTTCTGACAGTGTCTCTGCGGAGCTGCGCCAGCTGATG 360  
 QY 361 TACGTGGGCGCCGAGAGAGCCCTGTTGCTTGCCTTACAGATGAGAGCCCTGAGCTGCA 420  
 DB 361 TACGTGGGCGCCGAGAGAGCCCTGTTGCTTGCCTTACAGATGAGAGCCCTGAGCTGCA 420  
 QY 421 GCGATCTCTGGGAGAGCCCGCTGAGAGAGAGAGTGTATATCAGAGAGAGAGAGAG 480  
 DB 421 GCGATCTCTGGGAGAGCCCGCTGAGAGAGAGAGTGTATATCAGAGAGAGAGAGAG 480  
 QY 481 AACGAGACCGAGTGTCTTCACTTCCGCTTCTGCGAGCCCTTACATGAGCTCCCACTTC 540  
 DB 481 AACGAGACCGAGTGTCTTCACTTCCGCTTCTGCGAGCCCTTACATGAGCTCCCACTTC 540  
 QY 541 TACGTCTGTGACCTAGCGCTTCCAGCCCAAGTGCACCTAGCTACAGCTGCTACCTTC 600  
 DB 541 TACGTCTGTGACCTAGCGCTTCCAGCCCAAGTGCACCTAGCTACAGCTGCTACCTTC 600  
 QY 601 ACTTTGGAGCANTGAGAGTGTGAAGTGGAGAGGCAAGTGTCCCTATGAGCCCACTAAG 660  
 DB 601 ACTTTGGAGCANTGAGAGTGTGAAGTGGAGAGGCAAGTGTCCCTATGAGCCCACTAAG 660  
 QY 661 GGCATGCTGGGCTTCTGTGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 720  
 DB 661 GGCATGCTGGGCTTCTGTGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 720  
 QY 721 GGCAGGAGAACCTTATCTGCTGTAACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 DB 721 GGCAGGAGAACCTTATCTGCTGTAACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 QY 781 CTGGCCTTTTGGCTCAACGACCTCATCTTGTAGGCTGTGCTATGATGATGAGTGTG 840  
 DB 781 CTGGCCTTTTGGCTCAACGACCTCATCTTGTAGGCTGTGCTATGATGATGAGTGTG 840  
 QY 841 GGC 843  
 DB 841 GGC 843

Search completed: August 19, 2003, 13:25:51  
 Job time : 444 secs









QY 1026 CCCGACTGCGAGCTTCTACTTTCACACAGCTGCGAGCGGATGACACACCTGCGAGGACACTC 1085  
 Db 920 ACCCGAGAGACTCCCTTCTTCTACTTTCACAGTGTGCGAGCTGTCAGCGGCGTGTGTCAGCCT 979  
 QY 1086 CTGGCAACAACACACCTTCTTTGGGGTTTTCAGACACAGTGGGGTGCATGATCTACTGTC 1145  
 Db 980 CCGGG 1039  
 QY 1146 GGGCATCTGTGAGTACAGAGTGGAGAGATCCAGCGGGTGTGAGGGGGGGGGGGGGGGGGGG 1205  
 Db 1040 GGGCTGTGCGGCTTGTGACCTGACAGAGTGGGAGCTGTGTTGAAGGGGGGGGGGGGGGGGG 1099  
 QY 1206 GTACCATGAGGAGAGCGGAGAGTGGGAGCGCTACAC---TGACCTGTACACAGCGCTGCG 1262  
 Db 1100 GAGAGAGTCCCGCGGAGTGCATCTGAGCGCGGGTGGGAGATCAGTGTGCTCGACCGCCG 1159  
 QY 1263 GCGTGGCTGTGCTATTAACAACTGCGATGCGGCGGCGGCGGCTACACAGCTGCTGAGAGCT 1322  
 Db 1160 GCGCGGGTGTGCTGCGAGCGCGGGGATGAGTACAACTGCTGCGAGCGCT-----T 1210  
 QY 1333 ACCCGACAACATCTCAACTTCTGTCAGAGAGCGCGGCTGATGAGAGAGAGAGAGAGAGAG 1382  
 Db 1211 GCGGATGAGATCTCAACTTGTGCAAGAGCGGCGGCTGTGATGAGAGAGAGAGAGAGAGAG 1270  
 QY 1383 TCGGTGAGAGCGCGCGCGCTGTGCTGTAAGA 1411  
 Db 1271 GCTGGGCGCATGCGCGCTGTGATCTGCGGA 1299

## RESULT 4

US-09-077-940A-1

; Sequence 1, Application US/09077940A  
 ; Patent No. 6576441  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMURA, Toru et al.  
 ; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME  
 ; FILE REFERENCE: 0020-4426P  
 ; CURRENT APPLICATION NUMBER: US/09/077,940A  
 ; CURRENT FILING DATE: 1998-06-05  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 3692  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; NAME/KEY: 5'UTR  
 ; LOCATION: (1)..(18)  
 ; OTHER INFORMATION:  
 ; NAME/KEY: CDS  
 ; LOCATION: (19)..(2682)  
 ; OTHER INFORMATION:  
 ; NAME/KEY: 3'UTR  
 ; LOCATION: (2683)..(3653)  
 ; OTHER INFORMATION:  
 ; NAME/KEY: polyA\_site  
 ; LOCATION: (3654)..(3692)  
 ; OTHER INFORMATION:  
 ; US-09-077-940A-1

Query Match 5.3%; Score 122; DB 4; Length 3692;

Best Local Similarity 51.0%; Pred. No. 6,7e-20;

Matches 479; Conservative 0; Mismatches 410; Indels 51; Gaps 6;

QY 478 AACACACAGACGAGTCTTCACTTCCGCTTCCGACGAGCTTACAAATGCTCCGAC 537  
 Db 382 AAGCAAGAGGATGATGTCGAACTTGTCAAGGTGCTCTCTCTGAGACAAATCCACG 441  
 QY 538 CTGACGCTGTGTGAGCTTACGCTTCCAGCCCAAGTGCACCTAGCTGCAACATGCTCAGC 597  
 Db 442 CTCCTGTGTGAGGCTCAATGATCAATCCATCTGTGCAATTAAGATGATGACACA 501  
 QY 598 TTCACTTTTGAGCATGAGACTTTTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCT 657

Db 502 CT---GCAGCTTCTTGAGAGACAACATCAGTGTGATGCGCGCTGCGCCCTACAGCCCAAG 558  
 QY 658 AAGGCGCATGTGCGGCTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 717  
 Db 559 CATGCCAATGTGCGCTCTTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 618  
 QY 718 CTGGGACAGGAAACCATTAATCTGCTGTAACATGAGGGGGGGGGGGGGGGGGGGGGGGGGGG 775  
 Db 619 CTAGCCATTCAGAGCTGTATCTACGTAAGCTTGGGAGCCGGGGGGGGGGGGGGGGGGGGGGGG 678  
 QY 776 -AGTACTGCGCTTTTGGCTCAACAGACCTTCACTTTTGAAGCTGTGCTGTGATGATGAG 834  
 Db 679 AAGCATGACTCCAAAGTGTGTTAAAGAGCATCTTTGTGATGAGCGT----- 725  
 QY 835 AGTGTGGGAGCTTACGCGGGGAGAGAGCAAGTCTACTTCTTCTTCAAGAGGAGGAGCA 894  
 Db 726 -----GAGTGGGAAAGCCAGCTGCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 768  
 QY 895 GTGAGTCCGACTGCTATGCGGAGAGAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954  
 Db 769 ATGAGATTAACTATCTGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828  
 QY 955 GATATGGGGGG---GCGACGACCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1011  
 Db 839 GATGTGGGCGGCTCCCGACAGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888  
 QY 1012 CTGGCATGCTGTGCGCGGAG 1071  
 Db 889 CTCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948  
 QY 1072 CTGACAGAGACCTCTGCGCAACACACACACTTCTTGGGGTTTTCAGACAGAGTGGGGT 1131  
 Db 949 GGTGTGTGAGGCTTGGGCGCGCTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008  
 QY 1132 GACATGATACCTGTGCGGCTATCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1191  
 Db 1009 AGCATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068  
 QY 1192 GGGCCCTATTAAGAGTACCATGAG 1248  
 Db 1069 GGGCGCTTCCGAG 1128  
 QY 1249 GTACCCAGCCCTGCGGCTGTGCTGTGATTAACAACAGTGGAGCGGCGGAGCTACAC 1308  
 Db 1129 GTACCCAGCGGCGGCGGCGGCTGTGCTGTGAGCGCGGAGTATGACG-----TACAA 1179  
 QY 1309 AGCTCCCTGAGAGCTACCGGAGCAACATCCGCAACTGCTGCAAGAACACCGCGTATGAG 1368  
 Db 1180 GCATCAATGCGCTTCTGAGAGAGATTCGCACTTGTAAAGACCCACCGAGTATGAGC 1239  
 QY 1369 GAGCAGGTGGGCTGTGAG 1408  
 Db 1240 GAGAGGCTGCTCCCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1279

## RESULT 5

US-09-308-179B-2

; Sequence 2, Application US/09308179B

; Patent No. 6436669

; GENERAL INFORMATION:

; APPLICANT: INAGAKI, Shinobu

; APPLICANT: FURUYAMA, Taisuo

; TITLE OF INVENTION: NOVEL SEMAPHORIN GENES (1)

; FILE REFERENCE: 0020-4562P

; CURRENT APPLICATION NUMBER: US/09/308,179B

; CURRENT FILING DATE: 1999-05-14

; PRIOR APPLICATION NUMBER: PCT/JP97/04111

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: JAPAN 321068/1996

; PRIOR FILING DATE: 1996-11-15

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 2
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(2898)
; OTHER INFORMATION: Strandedness: Double-stranded
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(2898)
; OTHER INFORMATION: any n - a, c, g, t, unknown, or other
US-09-308-1795-2

```

Query Match 5.0%; Score 115; DB 4; Length 2898;

Best Local Similarity 48.6%; Pred. No. 3e-18;

Matches 560; Conservative 0; Mismatches 545; Indels 48; Gaps 7;

```

QY 310 GGCATCCAGAGCTTCCTGACACTGACGCTGACGACCCACCTGGCTTGTACGTGGC 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 GGTATCTGTGATCTCCATACAAATGCTGATGATGATCAAGAGCGCTTGTGGGA 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 370 GCCCGAGAGGCCCTGTTGCTTCAGCATGGA---GGCCCTGGAGCTGCAAGAGGCGATC 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 GGCAGAGACCTTGTCTATTCCTGAACTTGGAAAGAGTCACTGACGGCTACAGAGATA 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 427 TCCTGGAGAGCCCCCGTGGAGAAAGACTAGTGTTCAGAAAGGAGAAACCAACAG 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 TACTGGCCGAGCAGACAGATMAAGGTAGAAATGCTATATGAAAG---AAAAGACGA 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 487 ACCGAGTCTTCAACTTCATCCGCTTCCTGACGCCCTACAAATGCTCCACCTGTACGTC 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 AATGAGTGTCCAAATATATCCGGGTTTGCATCACAACAGACACACTTGTGACC 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 TGTGGACACCAAGCCCTCCAGCCCAAGTGCACCTACCTCAACATG-----591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 TGTGTCTAGTGGAGCTTTTGTTCACACTGTGCTTCACAGAGTGGCGGACCATTCAGAG 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 592 ---CTGACCTTCACTTTGGAGATGAGAGTTTGAAGTGGAGGAGGAGTGTCCCTAT 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 GAACCCCTGTTTCACTCGGAGAGTCAACAGATCTGAGAGAGAGAGGAGCAATGTCTTT 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 649 GACCCAGCTAAGGCCCATGCTGGCTTCTTGTGAGTGGAGTGAAGCTGACCTGACCACTC 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 GACCCCAACTCTCTTCTTGTGCTCCAGCTAGTGGGAATGAGCTTGTGCTGACTTAC 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 709 AACACTTCTGGGAGCAGAACCATATATCTGCTGAATGGGCCCCACACTCATG 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 946 AGTGAATATTGGGGCAGAGACTGGCGATCTTCCGAGCATGGGGAATTAGGCCATATT 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 769 AAGACAGATAGCTGGCTTTGGCTC---AAGCACTCACTTTGAGGCTGTGCTAT 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1006 CGCAGTGAAGATGACATGAGCGGCTCTGAAAGAACCAAAATTTGTAAGTTCTATATG 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 826 GTACTGAGAGTGTGGGACGCTTCACGGGGGAGACAGCAAGGTCTACTTCTTCTTCAGG 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1066 ATTCTGATTAACGAAGAC-----CGAGATGACAAACAAATGTACTTTTCTTACT 1116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 886 GAGCGGCACTGAGAGTCCGACTGTATGCCGAGCAGAGTGGTCTGTGTGGCCCGCTGC 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1117 GAAGAGGCGCTGGAGGGGAGAAACAGCCACACGATCTACACCCGAGTGGGGGCGCTG 1176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 946 TGCAGAGGCGATATGGGGGCGACAGACCTTGACAGAGAGTGGAGCAGTTCCTAG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1177 TGGGTGAATGATGAGAGAGACAGAGATCTGTGTGAACAGTGGAGACTTTCCTTAA 1236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1006 GCGCGGCTGGAGTCTGCCCCGAAGTGG-----CAGCTCTACTTCAACAGCATG 1056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1237 GCGGCGCTGGTGTCTCAGTGGCGGGAATGAAATGCAACATCTTGTAGAGACTA 1296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1057 CAGCGCATGACACACCTCGCAGAGACACTCTGCGACAAACACACTTCTTGGGGTTT 1116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1297 GAGGATGTGTTTACTGCGCAGAGATCTTAAGAAATCCAGTATATTGGACTGTT 1356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1117 CAAGCAGTGGGGTGCATGTACTCTGCGCCATCTGTGAGTACAGATTGAAGATC 1176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1357 AATAGTACAGCAATATATTAGAGGCATGCTGATGTGTGATCATCATGTCAAGTATC 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1177 CAGCGGCTGTTTGAAGGCCCCCTATTAAGAGTACCAGAGAGCCAGAAAGTGGAGCCG 1236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1417 CGGAAAGCCCTTAATGATGCTATGCTATTAAGAGAGGCCCTGAATACACTGTGCTA 1476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1237 TACAGTACCCCTTACCCAGCCCTCGGCTGCTGCTATTAACAATGGCATCGGCCG 1296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1477 TATGAGAAAGAAAGTCCCTTACCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1297 CAGGCTACACCAAGCTCCCTGAGACTCCGACACATCTCTCAACTGCTGCAAGAACAC 1356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1537 AA---GTATGGAACCAACAAAGATTAACCCGATGAGCCCATCGGTTCCGAGAGATGAT 1593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1357 CCGCTGATGAGAGACAGAGTGGGCGCTCGGTGAGAGCCGCCCTGCTGTGAAGAGGC 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1594 CCTCTAATGTATGAGCCCATTAACCTGTTCATTAATAAACCAATACTGTAAACAGAT 1653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1417 ACCAAGCTTCAACC 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1654 GGAATATCAACC 1666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 6

US-08-121-713D-53

Sequence 53, Application US/08121713D

Patent No. 5639856

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

APPLICANT: Kolodkin, Alex L.

APPLICANT: Matthes, David R.

APPLICANT: Bentley, David R.

APPLICANT: O'Connor, Timothy

TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 Bush Street, Suite 3200

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/121,713D

FILING DATE: 13-SEP-1993

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

TELEX:

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

NAME/KEY: CDS

LOCATION: 16...2331

US-08-121-713D-53

Query Match 4.9%; Score 112.4; DB 1; Length 2601;  
 Best Local Similarity 51.4%; Pred. No. 1.2e-17;  
 Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

```

QY 358 CTGTACGTGGGGGCGCCGAGAGGCGCTGTTGGCTTCAGCATGAGAGCCCTGAGCTGCAG 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 220 CTGTATGTTGGAGCAAGAGATCAGATATTTTCATTCGACCTGTTAATATCAAGATTTT 279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 GGAGCATCTCTCTGGAGAGCGCCCGTGGAGAGAGACTGACTGTATCCAGAAAGGAG 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 280 CAAAGATTTGTGTGGCCAGTATCTTACACCAAGAGATGATGCAAGGGCTGGAGAA 339
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 AACAAACAGACCGAGTGTCTTCACTTCATCCGCTCTGAGACCCCTACATGCTCCAC 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 340 GACATCTCGAAGAGATGTCTTAATTTCAATCAAGTACTTAAAGCATATATACAGACTAC 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 CTGTACGTCTGTGGAGACTGACCTTCAGCCCAAGTGCAGCTAGCTACATGCTC--- 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 400 TTGTACGCTGTGGAGAGGGGCTTTTCATCCAAATTTGACCTACATTTGAAATTTGACAT 459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 -----ACCTTCATTGGAGCATGAGATGAGATTTGAAGATGGAGGAG 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 460 CATCTGAGAGCAATATTTTAAAGTGGAGATCAGATTTTGAAGAGCGCGTGGAGAG 519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 640 TGTCCCTATGACCCAGCTAAGGCCATGCTGCGCTTCTTGAGATGTGAGCTGACTCG 699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 520 AGTCCATATGACCCCTAGCTGCTGACAGCATCCCTTTAATATGAGATGAGAAATTAACCT 579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 700 GCCACATCAACAACTTCTGGGAGGAGAACCCATTAATCCGCTAATCACTGGGCGCCAC 759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 580 GGAATCTGACGCTGATTTTATGGGAGAGCTTGTCTATCTTCCAAATCTTGGGAGCCAC 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 760 CACTCCATGAGACAGAGTACCTG---CCTTTTGGCTCAAGCACTCACTTGTAGGC 816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 640 CACCCATCAGAGACAGAGATGATTCAGAGTGGCTCAATGATCCAAAGTTACTTACT 699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 817 TCTGCTATATGACTGAGAGTGTGGGAGCTTCAAGGGGAGAGAGACAGCAAGTCTACTTC 876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 700 GCCACCTCATCTCAGAGAGTGA-----CAATCCCTGAAGATGACAAACTATATCTTT 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 TTCTTCAGAGGCGGAGTGAATGCTGCTGCTGAGGAGAGTGTGGCTGTGTG 936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 751 TTCTTCCTGAAATATGCAATAGATGAGAGAACCTCTGGAAGAGCTACTACGCTGAATA 810
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 937 GCCCGTGTCTGAGAGGCGGATATGGGGGCGGACGAGCCCTGACAGAGAGTGAAGCAG 996
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 811 GGTGATATGCAAGATGACTTTTGGAGGGACAGAGAGTGTGTAATTAATGAGACACA 870
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 997 TTCTTCAGAGGCGGCTGAGCTGCTGTGCCC 1028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 871 TTCTTCAGAGCTGCTGTGATTTGCTCAGTGC 902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

# RESULT 7

US-08-835-268-53  
 ; Sequence 53, Application US/08835268  
 ; Patent No. 5807826  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goodman, Corey S.  
 ; APPLICANT: Kolodkin, Alex L.  
 ; APPLICANT: Bentley, David R.  
 ; APPLICANT: O'Connor, Timothy  
 ; TITLE OF INVENTION: The Semaphorin Gene Family  
 ; NUMBER OF SEQUENCES: 100  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 Bush Street, Suite 3200  
 ; CITY: San Francisco  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94104

## COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-835-268-53

```

Query Match 4.9%; Score 112.4; DB 1; Length 2601;  
 Best Local Similarity 51.4%; Pred. No. 1.2e-17;  
 Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

```

QY 358 CTGTACGTGGGGGCGCCGAGAGGCGCTGTTGGCTTCAGCATGAGAGCCCTGAGCTGCAG 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 220 CTGTATGTTGGAGCAAGAGATCAGATATTTTCATTCGACCTGTTAATATCAAGATTTT 279
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 418 GGAGCATCTCTCTGGAGAGCGCCCGTGGAGAGAGACTGACTGTATCCAGAAAGGAG 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 280 CAAAGATTTGTGTGGCCAGTATCTTACACCAAGAGATGAGTGAAGTGGGCTGGAGAA 339
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 AACAAACAGACCGAGTGTCTTCACTTCATCCGCTCTGAGACCCCTACATGCTCCAC 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 340 GACATCTCGAAGAGATGTGCTAATTTCAATCAAGTACTTAAAGCATATATACAGACTAC 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 CTGTACGTCTGTGGAGACTGACCTTCAGCCCAAGTGCAGCTAGCTAACAATGCTC--- 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 400 TTGTACGCTGTGGAGAGGGGCTTTTCATCCAAATTTGACCTACATTTGAAATTTGACAT 459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 -----ACCTTCATTGGAGCATGAGATGAGATTTGAAGATGGAGGAG 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 460 CATCTGAGAGCAATATTTTAAAGTGGAGATCAGATTTTGAAGAGCGCGTGGAGAG 519
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 640 TGTCCCTATGACCCAGCTAAGGCCATGCTGCGCTTCTTGAGATGTGAGCTGACTCG 699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 520 AGTCCATATGACCCCTAGCTGCTGACAGCATCCCTTTAATATGAGAGGAAATTAACCT 579
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 700 GCCACATCAGACAACTTCTTGGGAGGAGAACCCATTAATCCGCTAATCACTGGGCGCCAC 759
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 580 GGAATCTGACGCTGATTTTATGGGAGGAGCTTTGCTATTCGGAACCTTGGGAGCCAC 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 760 CACTCCATGAGACAGAGTACCTG---CCTTTTGGCTCAAGCACTCACTTGTAGGC 816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 640 CACCCATCAGAGACAGAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCAATAGT 699
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 817 TCTGCTATATGACTGAGAGTGTGGGAGCTTCAAGGGGAGAGAGACAGCAAGTCTACTTC 876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 700 GCCACCTCATCTCAGAGAGTGA-----CAATCCCTGAAGATGACAAACTATATCTTT 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 877 TTCTTCAGAGGCGGAGTGAATGCTGCTGCTGAGGAGAGTGTGGCTGTGTG 936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 751 TTTCTCGTGAATAATGATGAGAACACTCTGGAAAAGCTACTACCGTAGATA 810  
 QY 937 GCCGTGTCTGCAAGGCGATATGAGGGGCGACGACCTGACAGAGAAATGAGCAG 996  
 Db 811 GGTGAGATGATGCAAGATGACTTTGAGGGGCGACGAGATGCTGATTAATGACACAA 870  
 QY 997 TTCTGAGAGCGGCGCTGGCATGCTGCCCC 1028  
 Db 871 TTCTCAAAAGCTGCTGATTTGCTCAGTGCC 902

RESULT 8

US-09-060-692-53  
 Sequence 53, Application US/09060692  
 Patent No. 5935865

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.  
 APPLICANT: Kolodkin, Alex L.  
 APPLICANT: Mathes, David  
 APPLICANT: Bentley, David R.  
 APPLICANT: O'Connor, Timothy  
 TITLE OF INVENTION: The Semaphorin Gene Family  
 NUMBER OF SEQUENCES: 100  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 Bush Street, Suite 3200  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/060.692

FILING DATE:  
 CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/121.713

FILING DATE: 13-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.  
 REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415)343-4341

TELEFAX: (415) 343-4342

TELEX:  
 INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 16..2331

US-09-060-692-53

Query Match 4.98; Score 112.4; DB 2; Length 2601;  
 Best Local Similarity 51.48; Pred. No. 1.2e-17;

Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

QY 358 CTGATGCTGGGCGCGGAGAGCGCTGTTGCTTTCAGCATGAGCGCTGAGCTGCAA 417  
 Db 220 CTGATGCTGGGCGCGGAGAGCATATATTTTCATTCGACGCTGTTATATCAAGATTTT 279

QY 418 GGAGCGATCTCTCTGGGAGCGCGCTGGAGAGAGACTGATTCATTCAGAAAAGGAG 477

Db 280 CAAAAGATTGTGTCGCAATATCTTACACCAAGAGATGAATGCAAGTGGCTGCAAA 339  
 QY 478 AACACCAAGACGAGTCTTCAACTTATCCGCTTCTGAGCGCTCAATGCTCCAC 537  
 Db 340 GACATCTGAAAGAAATGTGTAATTTATCATAGGTACTTAAAGATATATACAGAC 399  
 QY 538 CTGATGCTGTGGCAGCTTACCGCTTCCAGGCCAAGTGCACCTTACGTAACATGCTC 594  
 Db 400 TTGTACGCTGTGGAGAGCGGGCTTTTCAATCAATTTGCACTTCAATTTGAATGACAT 459  
 QY 595 -----ACCTTCACTTTGAGATGAGAGTTGAAGTGAAGGAGGAG 639  
 Db 460 CATCTGAGCAATATTTTAACTGAGAACTCAATTTGAAAAGCGCGCTGGAG 519  
 QY 640 TGTCCTATGACCGAGCTTAAGGCGCATGCTGAGCTTCTGTGATGAGTGTACTG 699  
 Db 520 AGTCATATGACCTTAAGCTGTACAGCATCCCTTTAATAGTGAATTTATCT 579  
 QY 700 GCCACTCAAACAATCTCTGGGCGAGGAAACCCATTTCTGCTAAATGAGGCGCCAC 759  
 Db 580 GGAATGCACTGATTTATATGAGGCGAGACTTGTCTATCTCCGAACTTGGGCGAC 639  
 QY 760 CACTTCATGAAAGAGTACTG---CCTTTGGCTCAACGAACTCACTTGTAGG 816  
 Db 640 CACCAATCAGAGAGAGAGAGATGATTCAGGTGCTCAATGATCCAAAGTTGATAGT 699  
 QY 817 TCTGCTATGATGATGAGTGTGGGCGAGCTTCAAGGGGAGAGAGAGTGTACTG 876  
 Db 700 GCCACCTCATCTAGAGAGT-----CAATCTGAAGTGAAGTATATCTT 750  
 QY 877 TTTCTCAGAGAGCGGAGTGAAGTCCGACTTATCCGAGAGTGTGCTGCTGTG 936  
 Db 751 TTTCTCGTGAATAATGCAATGATGAGAGAACTCTGAAAAGTACTACGCTAGATA 810  
 QY 937 GCCGTGTCTGCAAGGCGATATGAGGGGCGACGACCTGCGAGAGAAATGAGCAG 996  
 Db 811 GGTGAGATGATGCAAGATGACTTTTGGAGGCGACAGAAAGTGTGTAATTAATGACACA 870

RESULT 9

US-08-833-391-53  
 Sequence 53, Application US/08833391  
 Patent No. 6013781

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.  
 APPLICANT: Kolodkin, Alex L.  
 APPLICANT: Mathes, David  
 APPLICANT: Bentley, David R.  
 APPLICANT: O'Connor, Timothy  
 TITLE OF INVENTION: The Semaphorin Gene Family  
 NUMBER OF SEQUENCES: 100  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 Bush Street, Suite 3200  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/833.391  
 FILING DATE:  
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/121.713

FILING DATE: 13-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman, Richard A.  
 REGISTRATION NUMBER: 36, 627  
 REFERENCE/DOCKET NUMBER: B94-002-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415)343-4341  
 TELEFAX: (415) 343-4342  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 53:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2601 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 16..2331  
 US-08-833-391-53

Query Match 4.9% Score 112.4; DB 3; Length 2601;

Best Local Similarity 51.4%; Pred. No. 1.2e-17;  
Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

358 CTGTACGTGGGCGCCGAGAGGCGCCCTGTTGCTTCAGCATGAGGCGCCCTGGAGCTGCA 417  
 220 CTGTATGTTGGAGCAAGAGATCATATTTTCATTCGACCTGTATATATCAAGATTTT 279  
 418 GGAGCATCTCTCTGGAGAGGCGCCCTGGAGAGAGATGATGATATCAGAAAGGGAAG 477  
 280 CAAAAGATTGTGTGGCCAGATATCTTACACACAGAGAGATGAGATGAGAGGCTGGAGAA 339  
 478 AACACAGAGAGAGATGCTTCAACTTCAATCCGCTTCCTGAGCGCCCTAGATGCTCCAG 537  
 340 GACATCTGGAAGAGATGCTTCAATTTCAATCAAGTACTTAAGGCAATATATGAGACTCAC 399  
 538 CTGTACGTCTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 594  
 400 TTGTAGGCTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 459  
 595 -----ACCTTCACTTTGGAGATGAGAGATGAGAGATGAGAGGAGGAGGAGGAGGAGG 639  
 460 CATCTGAGAGCAATATTTTAACTGAGAGAGATGAGAGATTTTGAAGAGGCGCTGGAGAG 519  
 640 TGTCCATATGAGCCAGCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 699  
 520 AGTCCATATGAGCCAGCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 579  
 700 GCCACACTCAGCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 759  
 580 GGAGCTGAGCTGATTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 639  
 760 CACTCTGAGAGAGAGAGTACCTGG---CCCTTGGCTCAAGAGAGGAGGAGGAGGAGGAGGAG 816  
 640 CACCCATATCAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 699  
 817 TCTGCTATGTAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 876  
 700 GCCCAGCTCATCTCAGAGAGTGA-----CAATCTGGAAGATGAGAGAGATATACCTTT 750  
 877 TTCTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 936  
 751 TTCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 810  
 937 GCCCGTGTCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 996  
 811 GGTCTGATATGCAAGATGAGCTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 870  
 997 TTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1028  
 871 TTCTCAAAAGGCTGCTGATTTTCTCAGTGGC 902

## RESULT 10

US-09-060-610-53

Sequence 53, Application US/09060610

Patent No. 6344544

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

APPLICANT: Kolodkin, Alex L.

APPLICANT: Mathes, David

APPLICANT: Bentley, David R.

APPLICANT: O'Connor, Timothy

TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 Bush Street, Suite 3200

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/060,610

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/835,268

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36, 627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)343-4341

TELEFAX: (415) 343-4342

TELEX:

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 16..2331

US-09-060-610-53

Query Match 4.9% Score 112.4; DB 4; Length 2601;

Best Local Similarity 51.4%; Pred. No. 1.2e-17;  
Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

358 CTGTACGTGGGCGCCGAGAGGCGCCCTGTTGCTTCAGCATGAGGCGCCCTGGAGCTGCA 417  
 220 CTGTATGTTGGAGCAAGAGATCATATTTTCATTCGACCTGTATATATCAAGATTTT 279  
 418 GGAGCATCTCTCTGGAGAGGCGCCCTGGAGAGAGATGATGATATCAGAAAGGGAAG 477  
 280 CAAAAGATTGTGTGGCCAGATATCTTACACACAGAGAGATGAGATGAGAGGCTGGAGAA 339  
 478 AACACAGAGAGAGATGCTTCAACTTCAATCCGCTTCCTGAGCGCCCTAGATGCTCCAG 537  
 340 GACATCTGGAAGAGATGCTTCAATTTCAATCAAGTACTTAAGGCAATATATGAGACTCAC 399  
 538 CTGTACGTCTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 594  
 400 TTGTAGGCTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 459  
 595 -----ACCTTCACTTTGGAGATGAGAGATGAGAGATGAGAGGAGGAGGAGGAGGAGG 639



```

Db      460 CATCTGAGAGCAATATTTTAACTGGAGACTCAATTTTAAAGCGCCCTGGGAG 519
      640 TGTCCCTATGACCCAGCTAAGGGCAGTGGCTTCTGTGGATGGTGAAGTACTG 699
      520 AGTCCATATGACCTTAAGCTGCTGACAGCAATCCCTTTTATATGAGAAATATACT 579
      700 GCCACACTCAACAATCTCTGCGGACGAGAACCCATTATCTGCTAATAGGGGCC 759
      580 GGAACCTGACAGCTATTTTATGAGGGGAGACTTGTATCTTCGAACTCTTGGGACAC 639
      760 CACTCATGAGACAGATGACTG---CCTTTGGCTCAGACACTCTTTGAGGC 816
      640 CACCCATGACGACAGCAGATGATCCAGGTGCTCATATCCAAAGTTCATTA 699
      817 TCTGCTATGACTGAGAGTGGGAGAGTTCACGGGGGACGACGAGAGTCTACT 876
      700 GCCCACTCATCTCAGAGAGTGA-----CAATCTGAAGTACAAAGTACTT 750
      877 TTCTCAGGAGCGGCGAGTGAAGTCCGACTGCTATGCCAGAGAGTGTGCTG 936
      751 TTCTCCGTGAATAATGAAATAGATGAGAACACTCTGGAAGAGTACTCAGCGTAGA 810
      937 GCCCGTCTGCAAGGCGCATATGCGGGGCGCAGGACCTCGAGAGAAAGTGGAC 996
      811 GGTCAATATGCAAGATGACTTGGAGGCGACAGAAAGTGTGAATTAATGACACA 870
      997 TTCTGAAGCGGCGCTGGCATGCTGCCCC 1028
      871 TTCTCAAGCTGCTGATTTCTCTCAGTGCC 902

```

RESULT 11  
PCT-US94-10151A-53

```

; Sequence 53, Application PC/TUS9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR HOBBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10151A
; FILING DATE: 13-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: EP-58750-PC/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331

```

PCT-US94-10151A-53

Query Match 4.9%; Score 112.4; DB 5; Length 2601;  
Best Local Similarity 51.4%; Pred. No. 1.2e-17;  
Matches 356; Conservative 0; Mismatches 306; Indels 30; Gaps 3;

```

      358 CTGTACGTGGGCGCCGAGAGGCGCTGTTGCTTCCAGATGAGAGCCCTGAGCGCAA 417
      220 CTGTATGTTGAGCAAGAGATCAATATTTCAATTCAGCTGGTTAATATCAAGATTT 279
      418 GGAGCATCTCTCGGGAGGCGCCCGTGGAAAGAAAGCTAGTATCCGAAAGGAG 477
      280 CAAAGATTGTGTGGCCAGATCTTACACCAAGAGAGATTAATCAAGTGGCTGAAA 339
      478 AACACAGACAGAGTGTCTTCAACTATCCGCTTCCAGGCGCTCAATAGCTCCAC 537
      340 GACATCTGAAAGATGTGTTATTTATCATAGTACTTAAGCATATATTCAGACTCAC 399
      538 CTGTACGTGTGGACACTACGCTTCCAGCCCAAGTCAACCTACGTCATCATGCTC 594
      400 TTGTACGCTGTGGAAGCGGGGCTTTTCATTCATTTGACACTACATTTGAATTTG 459
      595 -----ACCTTCACTTTGGAGCATGGAAGTTTGAAGATGGGAAGGCGAG 639
      460 CATCTGAGAGCAATATTTTAAAGCTGAGAGCACTCAATTTTGAAGCGCCCTGG 519
      640 TGTCCCTATGACCCAGCTAAGGGCAGTGGCTTCTGTGATGGTGAAGTGAAGTACT 699
      520 AGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTATATGAGAAATATACT 579
      700 GCCCACTCAACAATCTCTGCGGACGGAACCCATTATCTGCTAATAGGGGCC 759
      580 GGAACCTGACAGTATTTTATGAGGGCGAGACTTGTCTTCCGCAACTCTTGGGCA 639
      760 CACTCCATGAAAGAGAGTACTG---CCTTTGGCTCAAGCAAGTCACTTGTAGGC 816
      640 CACCCATGAGAGAGAGAGCATGATTCAGGCTCATGATCCAAAGTTCATTAGT 699
      817 TCTGCTATGACTGAGAGTGGGCGAGCTTCAAGGGGACGACGACAAAGTCTACT 876
      700 GCCCACTCATCTCAGAGAGTGA-----CAATCTGAAGTACAAAGTACTT 750
      877 TTCTCAGGAGCGGCGAGTGAAGTCCGACTGCTATGCCAGAGAGTGTGCTGCTG 936
      751 TTCTCCGTGAATAATGAAATAGATGAGAAACACTCTGGAAGAGTACTCAGCGTAGA 810
      937 GCCCGTCTGCAAGGCGCATATGCGGGGCGCAGGACCTCGAGAGAAAGTGGAC 996
      811 GGTCAATATGCAAGATGACTTGGAGGCGACAGAAAGTGTGAATTAATGACACA 870
      997 TTCTGAAGCGGCGCTGGCATGCTGCCCC 1028
      871 TTCTCAAGCTGCTGATTTCTCTCAGTGCC 902

```

RESULT 12

US-09-254-594-5  
Sequence 5, Application US/09254594  
Patent No. 6566094

```

; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN X
; FILE REFERENCE: 0020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent version 3.0
; SEQ ID NO 5
; LENGTH: 2790
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Tissue Type: Child Brain
; NAME/KEY: CDS
; LOCATION: (1)..(2790)
; OTHER INFORMATION: Identification Method: E
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Identification Method: P for resulting peptide
US-09-254-594-5

```

```

Query Match 4.8%; Score 108.8; DB 4; Length 2790;
Best Local Similarity 48.7%; Pred. No. 9e-17;
Matches 511; Conservative 0; Mismatches 492; Indels 47; Gaps 6;

```

```

QY 490 GAGTCTTCACATTCGCTTCCTGCGACCTCAATGCTCCACCTGCTACGTCGT 549
DB 358 GAGTCTTCACATTCGCTTCCTGCGACCTCAATGCTCCACCTGCTACGTCGT 417
QY 550 GGCACCTACGCTTCGACCCCAAGTGCACCTACGTCACATGCTCCTTCATTTGAG 609
DB 418 GGAACGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
QY 610 CATGAGAGTTTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTAAGGGCCATGCT 669
DB 475 GAGGTGAGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 534
QY 670 GGCCTTCCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 729
DB 535 GGCATCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 594
QY 730 CCCATTATCTGCGTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 786
DB 595 GCTGTAGTTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 654
QY 787 TTTTGGCTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846
DB 655 AAGTGTGCTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684
QY 847 TTCACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
DB 685 GCGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 744
QY 907 TGTCTATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
DB 745 GCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 804
QY 967 ---GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1023
DB 805 TCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 864
QY 1024 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1083
DB 865 GTCCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 924
QY 1084 TCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1143
DB 925 CTGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 984
QY 1144 TCGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1203
DB 985 TCTGCGCTGCTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1044
QY 1204 GAGTACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1263
DB 1045 GAGCA---GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1100
QY 1264 CCGTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1319
DB 1101 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1160
QY 1320 GCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1379

```

```

DB 1161 COTCCCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
QY 1380 GCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1439
DB 1221 ACCTGTACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1280
QY 1440 CGACCGGTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1499
DB 1281 TGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1340
QY 1500 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529
DB 1341 TGGACAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1370

```

```

RESULT 13
US-09-254-594-4
; Sequence 4, Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIMURA, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 0020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 3432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Tissue Type: Child Brain
; NAME/KEY: 5'UTR
; LOCATION: (1)..(187)
; OTHER INFORMATION: Identification Method: E
; NAME/KEY: misc.feature
; LOCATION: (188)..(2977)
; OTHER INFORMATION: CDS; Identification Method: E
; NAME/KEY: 3'UTR
; LOCATION: (2978)..(3407)
; OTHER INFORMATION: Identification Method: E
; NAME/KEY: polyA.signal
; LOCATION: (3408)..(3432)
; OTHER INFORMATION: Identification Method: E
US-09-254-594-4

```

```

Query Match 4.8%; Score 108.8; DB 4; Length 3432;
Best Local Similarity 48.7%; Pred. No. 9.7e-17;
Matches 511; Conservative 0; Mismatches 492; Indels 47; Gaps 6;

```

```

QY 490 GAGTCTTCACATTCGCTTCCTGCGACCTCAATGCTCCACCTGCTACGTCGT 549
DB 358 GAGTCTTCACATTCGCTTCCTGCGACCTCAATGCTCCACCTGCTACGTCGT 417
QY 550 GGCACCTACGCTTCGACCCCAAGTGCACCTACGTCACATGCTCCTTCATTTGAG 609
DB 418 GGAACGACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474
QY 610 CATGAGAGTTTGAAGATGGAAGGCAAGTGTCCCTATGACCCAGCTAAGGGCCATGCT 669
DB 475 GAGGTGAGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 534
QY 670 GGCCTTCCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 729
DB 535 GGCATCTTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 594
QY 730 CCCATTATCTGCGTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 786
DB 595 GCTGTAGTTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 654
QY 787 TTTTGGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846
DB 655 AAGTGTGCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684
QY 847 TTCACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
DB 685 GCGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 744
QY 907 TGTCTATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
DB 745 GCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 804
QY 967 ---GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1023
DB 805 TCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 864
QY 1024 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1083
DB 865 GTCCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 924
QY 1084 TCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1143
DB 925 CTGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 984
QY 1144 TCGGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1203
DB 985 TCTGCGCTGCTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1044
QY 1204 GAGTACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1263
DB 1045 GAGCA---GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1100
QY 1264 CCGTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1319
DB 1101 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1160
QY 1320 GCTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1379

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 14:58:57 ; Search time 418 Seconds  
(without alignments)  
12228.363 Million cell updates/sec

Title: US-10-002-050-13

Perfect score: 2284  
Sequence: 1 cggcccttcacactctctg.....tgtgtgtgtatgtgtgtgt 2284

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 308958

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published\_Applications\_NA\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2280	99.8	2284	13	US-10-003-152-13 Sequence 13, Appl
2	2280	99.8	2284	14	US-10-002-050-13 Sequence 13, Appl
3	2280	99.8	2284	14	US-10-002-304-13 Sequence 11, Appl
4	1887.4	82.6	2155	13	US-10-003-152-11 Sequence 11, Appl
5	1887.4	82.6	2155	14	US-10-002-050-11 Sequence 11, Appl
6	1887.4	82.6	2155	14	US-10-002-304-11 Sequence 11, Appl
7	1879.4	82.3	2156	13	US-10-003-152-21 Sequence 21, Appl
8	1879.4	82.3	2156	14	US-10-002-050-21 Sequence 21, Appl
9	1879.4	82.3	2156	14	US-10-002-304-21 Sequence 21, Appl
10	1761	77.1	3293	14	US-10-149-819-25 Sequence 25, Appl
11	1394	61.0	8095	10	US-09-989-920-73 Sequence 73, Appl
12	1351	59.2	3503	11	US-09-759-130B-336 Sequence 336, Appl
13	1351	59.2	3503	14	US-10-189-123-66 Sequence 66, Appl
14	1193.6	52.3	2811	11	US-09-759-130B-303 Sequence 303, Appl
15	1193.6	52.3	2811	14	US-10-189-123-33 Sequence 33, Appl
16	841.4	36.8	967	13	US-10-000-512-11 Sequence 11, Appl

17	463	20.3	729	11	US-09-759-130B-304	Sequence 304, App
18	463	20.3	729	14	US-10-189-123-34	Sequence 34, Appl
19	411	18.0	3781	11	US-09-946-374-252	Sequence 252, App
20	411	18.0	3781	12	US-10-015-387A-252	Sequence 252, App
21	411	18.0	3781	12	US-10-006-130A-252	Sequence 1, Appl
22	411	18.0	3781	12	US-10-199-672-453	Sequence 453, App
23	411	18.0	3781	12	US-10-241-220-1	Sequence 453, App
24	411	18.0	3781	13	US-10-052-586-453	Sequence 453, App
25	411	18.0	3781	14	US-10-174-590-453	Sequence 453, App
26	411	18.0	3781	14	US-10-176-758-453	Sequence 453, App
27	411	18.0	3781	14	US-10-175-737-453	Sequence 453, App
28	411	18.0	3781	14	US-10-173-706-453	Sequence 453, App
29	411	18.0	3781	14	US-10-175-738-453	Sequence 453, App
30	411	18.0	3781	14	US-10-175-752-453	Sequence 453, App
31	411	18.0	3781	14	US-10-176-482-453	Sequence 453, App
32	411	18.0	3781	14	US-10-176-757-453	Sequence 453, App
33	411	18.0	3781	14	US-10-176-913-453	Sequence 453, App
34	411	18.0	3781	14	US-10-180-552-453	Sequence 453, App
35	411	18.0	3781	14	US-10-180-557-453	Sequence 453, App
36	411	18.0	3781	14	US-10-173-700-453	Sequence 453, App
37	411	18.0	3781	14	US-10-174-572-453	Sequence 453, App
38	411	18.0	3781	14	US-10-174-579-453	Sequence 453, App
39	411	18.0	3781	14	US-10-174-582-453	Sequence 453, App
40	411	18.0	3781	14	US-10-174-588-453	Sequence 453, App
41	411	18.0	3781	14	US-10-175-739-453	Sequence 453, App
42	411	18.0	3781	14	US-10-175-740-453	Sequence 453, App
43	411	18.0	3781	14	US-10-175-743-453	Sequence 453, App
44	411	18.0	3781	14	US-10-176-488-453	Sequence 453, App
45	411	18.0	3781	14	US-10-176-492-453	Sequence 453, App

ALIGNMENTS

RESULT 1  
US-10-003-152-13  
Sequence 13, Application US/10003152  
Publication No. US20020151494A1  
GENERAL INFORMATION:  
APPLICANT: Shinketsu, Richard  
APPLICANT: Fernandes, Elma  
APPLICANT: Vernet, Corine  
APPLICANT: Yang, MeiJia  
APPLICANT: Boldog, Ferenc  
TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin  
FILE REFERENCE: 15966-554 Cura-54 CON-S12  
CURRENT APPLICATION NUMBER: US/10/003,152  
CURRENT FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: 09/604,286  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/140,584  
PRIOR FILING DATE: 1999-06-23  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 2284  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (166)..(1953)  
NAME/KEY: variation  
LOCATION: (1)..(2284)  
OTHER INFORMATION: N may be any nucleotide  
US-10-003-152-13

Query Match 99.8% Score 2280; DB 13; Length 2284;  
Best Local Similarity 100.0% Pred. No. 0;  
Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGCCCTTCACACTCCTGCCCTGATGTGGAACGGGGTTGGGTTCTGCAGGCT 60  
|||||

Db 1 CGGCCCTTCACACACTCTGCGCTGCTGATGTGGAAAGGGGTTGGGGTCTTCGACAGGCT 60  
Oy 61 ATTGTCTGCGCTGGGGGAAGGGGACAGGCGGGGACCGGACCTCGCTTCGACAGCGGCGCC 120  
Db 62 ATTGTCTGCGCTGGGGGAAGGGGACAGGCGGGGACCGGACCTCGCTTCGACAGCGGCGCC 120  
Oy 121 ACCAGCAGAGAGCTGGGCTGAGAGCTCAGAGCCGGGGGCTGGCCATGAGCCGACACTGG 180  
Db 121 ACCAGCAGAGAGCTGGGCTGAGAGCTCAGAGCCGGGGGCTGGCCATGAGCCGACACTGG 180  
Oy 181 GCTGTCTGCTGTGGCAGCAGAGGCTGGGGGCTGGGCTATGGGGGCTGAGGTGGTGG 240  
Db 181 GCTGTCTGCTGTGGCAGCAGAGGCTGTGGGGGCTGGGCTATGGGGGCTGAGGTGGTGG 240  
Oy 241 AACCTGTGGCGGCTAAGACAGTGTCTTGGGGAGCTGGCCAGCGATGATACGGCGGTTTC 300  
Db 241 AACCTGTGGCGGCTAAGACAGTGTCTTGGGGAGCTGGCCAGCGATGATACGGCGGTTTC 300  
Oy 301 TCCGAGACCGGCTATCCAGAGATTTCTCTGACACTGAGAGCTGAGAGGCCCTGAGCTGCA 360  
Db 301 TCCGAGACCGGCTATCCAGAGATTTCTCTGACACTGAGAGCTGAGAGGCCCTGAGCTGCA 360  
Oy 361 TACGTGGGCGGCGGAGAGAGGCTGTTTGCCTTCAGCATGAGAGGCCCTGAGCTGCAAGGA 420  
Db 361 TACGTGGGCGGCGGAGAGAGGCTGTTTGCCTTCAGCATGAGAGGCCCTGAGCTGCAAGGA 420  
Oy 421 GCGATCTCTCTGGAGAGGCGGCGGTGGAGAGAGACTGATGTATCCAGAAAGGAGAAC 480  
Db 421 GCGATCTCTCTGGAGAGGCGGCGGTGGAGAGAGACTGATGTATCCAGAAAGGAGAAC 480  
Oy 481 AACGAGACCGAGTGTCACTCACTGCTTCGAGCGGCTCAATGGCCCTCCACCTG 540  
Db 481 AACGAGACCGAGTGTCACTCACTGCTTCGAGCGGCTCAATGGCCCTCCACCTG 540  
Oy 541 TACGCTGTGGACCTTACGCTTCAGGCCCAAGTGCACACTGATCAACATGCTCACTTC 600  
Db 541 TACGCTGTGGACCTTACGCTTCAGGCCCAAGTGCACACTGATCAACATGCTCACTTC 600  
Oy 601 ACTTGGAGCATGAGAGATTTGAAAGATGGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG 660  
Db 601 ACTTGGAGCATGAGAGATTTGAAAGATGGGAAAGGCAAGTGTCCCTATGACCCAGCTAAG 660  
Oy 661 GACCATGTGGGCTCTTGTGGATGGTGAAGTGTACTGAGGCAACCTCAACAACTCTG 720  
Db 661 GACCATGTGGGCTCTTGTGGATGGTGAAGTGTACTGAGGCAACCTCAACAACTCTG 720  
Oy 721 GGCAGCGAACCCATTATCTGCTGATTAACATGGGGGCGCCACACTGCATGAAGACAGATAC 780  
Db 721 GGCAGCGAACCCATTATCTGCTGATTAACATGGGGGCGCCACACTGCATGAAGACAGATAC 780  
Oy 781 CTGGGCTTTTGGCTCAGCAAGCACTTCATTTGAGGCTGTGCCATATGACCTGAGAGTGG 840  
Db 781 CTGGGCTTTTGGCTCAGCAAGCACTTCATTTGAGGCTGTGCCATATGACCTGAGAGTGG 840  
Oy 841 GGCAGCTTCACAGGGGAGAGAGAGAGTGTACTTCTTCTTCAAGGAGCGGCACTGGAG 900  
Db 841 GGCAGCTTCACAGGGGAGAGAGAGAGTGTACTTCTTCTTCAAGGAGCGGCACTGGAG 900  
Oy 901 TCCGACTCTATGCGAGCAGAGTGTGGCTGTGGGCGGCTGTCTGCAAGGGCGCATATG 960  
Db 901 TCCGACTCTATGCGAGCAGAGTGTGGCTGTGGGCGGCTGTCTGCAAGGGCGCATATG 960  
Oy 961 GGGGGCGGAGGAGGAGGAG 1020  
Db 961 GGGGGCGGAGGAGGAGGAG 1020  
Oy 1021 TCTGCGCCGAACTGGCACTCTACTTCAACAGAGCTGAGGGGATGACACCTGAGAGAC 1080  
Db 1021 TCTGCGCCGAACTGGCACTCTACTTCAACAGAGCTGAGGGGATGACACCTGAGAGAC 1080  
Oy 1081 ACCTCTCTGAGCAACACCACTCTTCTTGGGGTTTTTCAAGCAGACAGTGGGATGATAC 1140  
Db 1081 ACCTCTCTGAGCAACACCACTCTTCTTGGGGTTTTTCAAGCAGACAGTGGGATGATAC 1140

Oy 1141 CTGTGCGCATCTGTGATACAGTGGAGAGATCCAGCGGGTGTGGAGGCGCCCTAT 1200  
Db 1141 CTGTGCGCATCTGTGATACAGTGGAGAGATCCAGCGGGTGTGGAGGCGCCCTAT 1200  
Oy 1201 AAGAGATCAATGAG 1260  
Db 1201 AAGAGATCAATGAG 1260  
Oy 1261 CGGCTGTGCTGTGATTAACAACTGGGCTGGGCGGCTGACAGGCTGCTGGAG 1320  
Db 1261 CGGCTGTGCTGTGATTAACAACTGGGCTGGGCGGCTGACAGGCTGCTGGAG 1320  
Oy 1321 CTACCGCAACATCTCACTTGTCAAGAGACCCGCTGATGAGAGAGAGAGAGAGAGAG 1380  
Db 1321 CTACCGCAACATCTCACTTGTCAAGAGACCCGCTGATGAGAGAGAGAGAGAGAGAG 1380  
Oy 1381 CCTGGTGGAGCGGCGGCGGCTGCTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Db 1381 CCTGGTGGAGCGGCGGCGGCTGCTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Oy 1441 GACCGGGTTACAGAGACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
Db 1441 GACCGGGTTACAGAGACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
Oy 1501 GCGTGGCTGTCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Db 1501 GCGTGGCTGTCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Oy 1561 CTGTTTGAACAG 1620  
Db 1561 CTGTTTGAACAG 1620  
Oy 1621 GCGGCTCCCGCTCAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
Db 1621 GCGGCTCCCGCTCAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
Oy 1681 TGTGCAAGCTGTGCTGTGCTGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Db 1681 TGTGCAAGCTGTGCTGTGCTGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Oy 1741 TGTGAGGCGGAG 1800  
Db 1741 TGTGAGGCGGAG 1800  
Oy 1801 ACTTCAGGCAATTTGCAACCTCGGTGGAGTAAAGATACAGTACAGGCCACTNCCAAAAC 1860  
Db 1801 ACTTCAGGCAATTTGCAACCTCGGTGGAGTAAAGATACAGTACAGGCCACTNCCAAAAC 1860  
Oy 1861 ATCAGGCTGTGGGCGGAG 1920  
Db 1861 ATCAGGCTGTGGGCGGAG 1920  
Oy 1921 CTGCGGAGCTCAACCCGAGAGAGATCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
Db 1921 CTGCGGAGCTCAACCCGAGAGAGATCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
Oy 1981 ANAGGCTGGAGAGTGTACTCTTACTTTTGGACAGGACACAGCTATTCAGGGACATGGC 2040  
Db 1981 ANAGGCTGGAGAGTGTACTCTTACTTTTGGACAGGACACAGCTATTCAGGGACATGGC 2040  
Oy 2041 ACGGCACTGTCTGTGTGGAGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
Db 2041 ACGGCACTGTCTGTGTGGAGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
Oy 2101 GCTGTGTCTCAG 2160  
Db 2101 GCTGTGTCTCAG 2160  
Oy 2161 CATCTTCTCTCTCTGTGAATCAGAGACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
Db 2161 CATCTTCTCTCTCTGTGAATCAGAGACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220

QY 2221 GCAGAACTTMAAGATGTGTTTGTATTTGACATGTCGTTTGTCTGTGTATGTC 2280  
DB 2221 GCAGAACTTMAAGATGTGTTTGTATTTGACATGTCGTTTGTCTGTGTATGTC 2280  
QY 2281 TGTG 2284  
DB 2281 TGTG 2284

RESULT 2

US-10-002-050-13  
Sequence 13, Application US/10002050  
Publication No. US20030032095A1  
GENERAL INFORMATION:  
APPLICANT: Shimkets, Richard  
APPLICANT: Fernandes, Elma  
APPLICANT: Vernet, Corline  
APPLICANT: Yang, Meljia  
APPLICANT: Boldog, Ferenc  
APPLICANT: Heitmann, John  
TITLE OF INVENTION: No. US20030032095A1e1 Nucleic Acid Sequences Encoding Human Semaphorin 4D  
FILE REFERENCE: 15966-554 Cura-54 CON-S14  
CURRENT APPLICATION NUMBER: US/10/002, 050  
CURRENT FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: 09/604, 286  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/140, 584  
PRIOR FILING DATE: 1999-06-23  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 13  
LENGTH: 2284  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (166)..(1953)  
NAME/KEY: variation  
LOCATION: (1)..(2284)  
OTHER INFORMATION: N may be any nucleotide  
US-10-002-050-13

Query Match 99.8%; Score 2280; DB 14; Length 2284;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCGCTTCTCAACACCTCCCTGCTGATGTGGAACGGGGTTTGGGGTTCTGACAGGCT 60  
DB 1 CGGCGCTTCTCAACACCTCCCTGCTGATGTGGAACGGGGTTTGGGGTTCTGACAGGCT 60  
QY 61 ATTGTGCGCTGCGGGAAGGGGACAGCGCGGACCGGGAACCTCCGCTCGACGCGGCGCG 120  
DB 61 ATTGTGCGCTGCGGGAAGGGGACAGCGCGGACCGGGAACCTCCGCTCGACGCGGCGCG 120  
QY 121 ACCAGACGACAGCTGGCTGAAAGCTCAGAGCCGGGGCGTGGCCATGGCCCAACACTGG 180  
DB 121 ACCAGACGACAGCTGGCTGAAAGCTCAGAGCCGGGGCGTGGCCATGGCCCAACACTGG 180  
QY 121 ACCAGACGACAGCTGGCTGAAAGCTCAGAGCCGGGGCGTGGCCATGGCCCAACACTGG 180  
DB 121 ACCAGACGACAGCTGGCTGAAAGCTCAGAGCCGGGGCGTGGCCATGGCCCAACACTGG 180  
QY 181 GCTGTGCTGCTGCGACAGCAAGCTGTGGGGCTGGGGCACTGGGGCTGAGGTTGGTGG 240  
DB 181 GCTGTGCTGCTGCGACAGCAAGCTGTGGGGCTGGGGCACTGGGGCTGAGGTTGGTGG 240  
QY 241 AACCTTGTGCGCGGTAAAGACAGTGTCTGTGGGAGCTGGCCACGGTAGTACGGGGTTC 300  
DB 241 AACCTTGTGCGCGGTAAAGACAGTGTCTGTGGGAGCTGGCCACGGTAGTACGGGGTTC 300  
QY 301 TCCCAACCGGCGCATCCAGAGACTCTGACACTACAGCTGACGAGGCCCACTGGGCTTCTG 360  
DB 301 TCCCAACCGGCGCATCCAGAGACTCTGACACTACAGCTGACGAGGCCCACTGGGCTTCTG 360  
QY 361 TACGTGGGGCGCGGAGAGGCGCTGTGCTGAGATGAGAGGCGCTGAGAGCGCAAGGA 420  
DB 361 TACGTGGGGCGCGGAGAGGCGCTGTGCTGAGATGAGAGGCGCTGAGAGCGCAAGGA 420

QY 421 GCGATCTCTGSGAGGCCCCCTGGAGAAAGACTGAGTGTATCCAGAAAGGAAAGAC 480  
DB 421 GCGATCTCTGSGAGGCCCCCTGGAGAAAGACTGAGTGTATCCAGAAAGGAAAGAC 480  
QY 481 AACCAAGCGAGTGTCAACCTTCCAGTCCGCTGACGCTTACAAATGCTCCACCTG 540  
DB 481 AACCAAGCGAGTGTCAACCTTCCAGTCCGCTGACGCTTACAAATGCTCCACCTG 540  
QY 541 TAGCTGTGSGACATACGCTTCCAGCCCAATGCAACCTGACATGCTGACCTTC 600  
DB 541 TAGCTGTGSGACATACGCTTCCAGCCCAATGCAACCTGACATGCTGACCTTC 600  
QY 601 ACTTTGAGAGATGAGAGTGTGAAGATGGAAGGCAAGTGTCCATGACCCAGCTAAG 660  
DB 601 ACTTTGAGAGATGAGAGTGTGAAGATGGAAGGCAAGTGTCCATGACCCAGCTAAG 660  
QY 661 GGCATGCTGCGCTTCTGTGTGATGTTGAGCTGTACTGGCCACACTCAACACTTCCTG 720  
DB 661 GGCATGCTGCGCTTCTGTGTGATGTTGAGCTGTACTGGCCACACTCAACACTTCCTG 720  
QY 721 GGCAGGAAACCAATTAATCCGCTTAACATAGGCGCCACACTCCATGAAAGAGATAC 780  
DB 721 GGCAGGAAACCAATTAATCCGCTTAACATAGGCGCCACACTCCATGAAAGAGATAC 780  
QY 781 CTGGCTTTTGGCTCAACGAACTTGTAGGCTGTCTATGATGATGAGAGTGTG 840  
DB 781 CTGGCTTTTGGCTCAACGAACTTGTAGGCTGTCTATGATGATGAGAGTGTG 840  
QY 841 GGCAGCTTCAAGGCGGAGACGACAAAGTCTACTTCTTCAAGGAGCGGCGAGTGTG 900  
DB 841 GGCAGCTTCAAGGCGGAGACGACAAAGTCTACTTCTTCAAGGAGCGGCGAGTGTG 900  
QY 901 TCCGACTGCTATCCGAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
DB 901 TCCGACTGCTATCCGAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
QY 961 GGGGCGGACGAGACCTGACAGAAAGTGTGAGCAAGCTTCTGAAAGGCGGCTGGCATG 1020  
DB 961 GGGGCGGACGAGACCTGACAGAAAGTGTGAGCAAGCTTCTGAAAGGCGGCTGGCATG 1020  
QY 1021 TCTGCCCCGAACTGGAGCTCTACTTCAACCACTGAGGCGATGACACCTGCAAGAG 1080  
DB 1021 TCTGCCCCGAACTGGAGCTCTACTTCAACCACTGAGGCGATGACACCTGCAAGAG 1080  
QY 1081 ACCTGCTGCAACCAACCACTTCTTGGGCTTTTCAACCAAGTGGGTGACATGTAC 1140  
DB 1081 ACCTGCTGCAACCAACCACTTCTTGGGCTTTTCAACCAAGTGGGTGACATGTAC 1140  
QY 1141 CTGTGCGCATCTGTGAGTACAGTGTGAAAGATGACAGGCGTGTGTGAGGCGCCCTAT 1200  
DB 1141 CTGTGCGCATCTGTGAGTACAGTGTGAAAGATGACAGGCGTGTGTGAGGCGCCCTAT 1200  
QY 1201 AAGGATGATCATGAGAGAGCCGAGAAAGTGGAGCCGCTACACTGACCTGTACCAAGCCT 1260  
DB 1201 AAGGATGATCATGAGAGAGCCGAGAAAGTGGAGCCGCTACACTGACCTGTACCAAGCCT 1260  
QY 1261 CGGCTGCGCTGCTGATTAACAAGTGGCACTGGCGCAGCGGCTACACCACTCCCTGGAG 1320  
DB 1261 CGGCTGCGCTGCTGATTAACAAGTGGCACTGGCGCAGCGGCTACACCACTCCCTGGAG 1320  
QY 1321 CTACCCGACAAATCTCAACTTCTGTCAGAAAGACCCGCTGATGAGAGGAGAGTGGGG 1380  
DB 1321 CTACCCGACAAATCTCAACTTCTGTCAGAAAGACCCGCTGATGAGAGGAGAGTGGGG 1380  
QY 1381 CTTGCGTGAAGCCGCCCTGCTGTGTGAAGAGGCAACCACTTCAACCACTGCTGGTGGC 1440  
DB 1381 CTTGCGTGAAGCCGCCCTGCTGTGTGAAGAGGCAACCACTTCAACCACTGCTGGTGGC 1440  
QY 1441 GACCGGGTTAAGAGACTGTATGAGGACCACTTATACAGTGTGTTCATTTGGACAGAGAC 1500  
DB 1441 GACCGGGTTAAGAGACTGTATGAGGACCACTTATACAGTGTGTTCATTTGGACAGAGAC 1500

```

OY 1501 GCGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGCTTACCTGATTTGAGAGCTGCAG 1560
DB 1501 GCGTGGCTGCTCAAGGCTGTGAGCTGGGGCCCTGGCTTACCTGATTTGAGAGCTGCAG 1560
OY 1561 CTGTTGACGAGAGCCCATGAGAAAGCTGGTGTATCTGACAGAGCAAGCTGCTCTTT 1620
DB 1561 CTGTTGACGAGAGCCCATGAGAAAGCTGGTGTATCTGACAGAGCAAGCTGCTCTTT 1620
OY 1621 GCCGGCTCCGCTCTGACCTGGTGCAGCTGCCCGCTGGCCGACTGCATTAAGTATGCTTC 1680
DB 1621 GCCGGCTCCGCTCTGACCTGGTGCAGCTGCCCGCTGGCCGACTGCATTAAGTATGCTTC 1680
OY 1681 TGTGAGAGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
DB 1681 TGTGAGAGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
OY 1741 TGTGAGAGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
DB 1741 TGTGAGAGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
OY 1801 ACTTACAGCATTTTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 1801 ACTTACAGCATTTTGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
OY 1861 ATTCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 1861 ATTCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
OY 1921 CTGCGCCGACTCAACCCCGAGAGTCACTATGAGAGGAAACCCCGCTGCGCGG 1980
DB 1921 CTGCGCCGACTCAACCCCGAGAGTCACTATGAGAGGAAACCCCGCTGCGCGG 1980
OY 1981 ANAGCTGTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
DB 1981 ANAGCTGTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
OY 2041 ACGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 2041 ACGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
OY 2101 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 2101 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
OY 2161 CATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
DB 2161 CATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
OY 2221 GCAGAGTTNAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
DB 2221 GCAGAGTTNAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
OY 2281 TGTG 2284
DB 2281 TGTG 2284

```

```

RESULT 3
us-10-002-304-13
: Sequence 13, Application US/10002304
: Publication No. US20030036185A1
: GENERAL INFORMATION:
: APPLICANT: Shimketa, Richard
: APPLICANT: Fernandes, Elma
: APPLICANT: Vernet, Corine
: APPLICANT: Yang, Melija
: APPLICANT: Boidoc, Ferenc
: APPLICANT: Herrmann, John
: TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
: FILE REFERENCE: 15966-554 Cura-54 CON-58
: CURRENT APPLICATION NUMBER: US/10/002,304
: PRIORITY FILING DATE: 2001-11-02
: PRIOR APPLICATION NUMBER: 09/604,286

```

```

: PRIOR FILING DATE: 2000-06-22
: PRIOR APPLICATION NUMBER: 60/140,584
: PRIOR FILING DATE: 1999-06-23
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 13
: LENGTH: 2284
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (166)..(1953)
: NAME/KEY: Variation
: LOCATION: (1)..(2284)
: OTHER INFORMATION: N may be any nucleotide
US-10-002-304-13

```

```

Query Match 99.8%; Score 2280; DB 14; Length 2284;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 CGGCCCTCTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
DB 1 CGGCCCTCTCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
OY 61 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
OY 121 ACCAGCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 121 ACCAGCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
OY 181 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
OY 241 AACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 AACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
OY 301 TCCGAGACGGGATCCAGAGACTTCTGACACTGACAGCTGACAGGAGCCCACTGAGGCTTCTG 360
DB 301 TCCGAGACGGGATCCAGAGACTTCTGACACTGACAGCTGACAGGAGCCCACTGAGGCTTCTG 360
OY 361 TACGTGGCGCCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 TACGTGGCGCCGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
OY 421 GCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
OY 481 AACGAGACGAGTGTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 540
DB 481 AACGAGACGAGTGTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 540
OY 541 TACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 TACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
OY 601 ACTTGGAGCATGAGAGATTGAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 601 ACTTGGAGCATGAGAGATTGAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
OY 661 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
OY 721 GGCAGGAAACCATTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 GGCAGGAAACCATTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

```



QY 781 CTGGCTTTTGGCTCAACGAACTCACTTTGTAGGCTCTGCTATGTACTGAGAGTGTG 840  
DB 781 CTGGCTTTTGGCTCAACGAACTCACTTTGTAGGCTCTGCTATGTACTGAGAGTGTG 840  
QY 841 GGGAGCTTCAGGGGGGAGGAGCAAGGCTACTGCTCTTCAGGGAGGGGAGAGTGGAG 900  
DB 841 GGGAGCTTCAGGGGGGAGGAGCAAGGCTACTGCTCTTCAGGGAGGGGAGAGTGGAG 900  
QY 901 TCCGACTGTATGTCGAGAGAGTGTGTGCTGTGTGGCCCGGTGTCTCAAGGGGAGATG 960  
DB 901 TCCGACTGTATGTCGAGAGAGTGTGTGCTGTGTGGCCCGGTGTCTCAAGGGGAGATG 960  
QY 961 GGGGGGAGCAGGAACTCTGAGAGAGAGTGGACCACTCTCTGAAGGCGCGGTGGAGTC 1020  
DB 961 GGGGGGAGCAGGAACTCTGAGAGAGAGTGGACCACTCTCTGAAGGCGCGGTGGAGTC 1020  
QY 1021 TCGGCCCCGAAACGAGAGTCTACTTCAACCACTGTAGAGCGATGCAACCCCTGAGAG 1080  
DB 1021 TCGGCCCCGAAACGAGAGTCTACTTCAACCACTGTAGAGCGATGCAACCCCTGAGAG 1080  
QY 1081 ACCTCTGAGCAACACCACTCTTTTGGGTTTTTCAAGCAGAGTGGGGTGAATGTAC 1140  
DB 1081 ACCTCTGAGCAACACCACTCTTTTGGGTTTTTCAAGCAGAGTGGGGTGAATGTAC 1140  
QY 1141 CTGTGCGCCATCTGTGTAGTACCAAGTGGAGAGATCCAGCGGGTGTGTGAAGGCGCTAT 1200  
DB 1141 CTGTGCGCCATCTGTGTAGTACCAAGTGGAGAGATCCAGCGGGTGTGTGAAGGCGCTAT 1200  
QY 1201 AAGGAGTACCATGAGAGAGGCGGAAAGTGGAGCCGCTACACTGACCCCTGTATCCAGCCCT 1260  
DB 1201 AAGGAGTACCATGAGAGAGGCGGAAAGTGGAGCCGCTACACTGACCCCTGTATCCAGCCCT 1260  
QY 1261 CGGCTGAGCTCTGTATTAACAAGTGGAGCGGAGGCTGACACAGCTCCCTGGAG 1320  
DB 1261 CGGCTGAGCTCTGTATTAACAAGTGGAGCGGAGGCTGACACAGCTCCCTGGAG 1320  
QY 1321 CTACCCGAGCAACATCTCTCAACTCTGTCAAGAGCAGCCGCTGATGAGAGAGAGTGGGG 1380  
DB 1321 CTACCCGAGCAACATCTCTCAACTCTGTCAAGAGCAGCCGCTGATGAGAGAGAGTGGGG 1380  
QY 1381 CCTGCGTGGAGCGGCGGCTGCTGTGAGAGAGGAGGAGCAACCTCACCCAGCTGGTGGCC 1440  
DB 1381 CCTGCGTGGAGCGGCGGCTGCTGTGAGAGAGGAGGAGCAACCTCACCCAGCTGGTGGCC 1440  
QY 1441 GACCGGGTTCAGAGACTTGTAGAGCCACTATACAGTGTGTTCATTTGGCAGAGAGAG 1500  
DB 1441 GACCGGGTTCAGAGACTTGTAGAGCCACTATACAGTGTGTTCATTTGGCAGAGAGAG 1500  
QY 1501 GGGTGGCTCTCAAGGCTGTGAGGCGGCGCTGGGTTCACTGATTTGAGAGAGCTGAG 1560  
DB 1501 GGGTGGCTCTCAAGGCTGTGAGGCGGCGCTGGGTTCACTGATTTGAGAGAGCTGAG 1560  
QY 1561 CTGTTTGAACGAGAGCCATGAGAAAGCTGTGTATCTCAAGAGCAAGAGCTCTCTTT 1620  
DB 1561 CTGTTTGAACGAGAGCCATGAGAAAGCTGTGTGTATCTCAAGAGCAAGAGCTCTCTTT 1620  
QY 1621 GCGGCTCCCGCTCTCAAGTGTGAGCTGCGGTGGCCGAGTCAATTAAGTATAGCTCC 1680  
DB 1621 GCGGCTCCCGCTCTCAAGTGTGAGCTGCGGTGGCCGAGTCAATTAAGTATAGCTCC 1680  
QY 1681 TGTGACAGACTGTGTCTGCGCGGAGCCCTATTTGCGCTGAGAGCTCAACAGGCGC 1740  
DB 1681 TGTGACAGACTGTGTCTGCGCGGAGCCCTATTTGCGCTGAGAGCTCAACAGGCGC 1740  
QY 1741 TGTGTGGCGGTGGTGGCCACTTGTATCTTACTGTATCAAGATGTGATGAGCTCGGAG 1800  
DB 1741 TGTGTGGCGGTGGTGGCCACTTGTATCTTACTGTATCAAGATGTGATGAGCTCGGAG 1800  
QY 1801 ACTTCAGAGATTTGCAACCTCTCGTGGAGATAGATACAGGCGCCACTNCCCAAAAC 1860  
DB 1801 ACTTCAGAGATTTGCAACCTCTCGTGGAGATAGATACAGGCGCCACTNCCCAAAAC 1860  
QY 1861 ATCAAGGTGGTGGCGGAGCAAGACTGTGTCTGCGCTGCAACCTCTCTCCCAACTTGGC 1920

DB 1861 ATCAAGGTGGTGGCGGAGCAAGACTGTGTCTGCGCTGCAACCTCTCTCCCAACTTGGC 1920  
QY 1921 CTGCGCGAGTCCAAACCCGAGAGTCACTAGTATGAGGGGAGACCCCAACCGCGTGGCGG 1980  
DB 1921 CTGCGCGAGTCCAAACCCGAGAGTCACTAGTATGAGGGGAGACCCCAACCGCGTGGCGG 1980  
QY 1981 ANAGCGTGGAGAGTGTAGTCTACTTGTTCACAGGCAAGAGTATCTCAGAGGAGATGCG 2040  
DB 1981 ANAGCGTGGAGAGTGTAGTCTACTTGTTCACAGGCAAGAGTATCTCAGAGGAGATGCG 2040  
QY 2041 ACGGCGACCTGTCTGTGTGGAGACAGATCTGCCAGACCCCAACCGCGCATGAGACT 2100  
DB 2041 ACGGCGACCTGTCTGTGTGGAGACAGATCTGCCAGACCCCAACCGCGCATGAGACT 2100  
QY 2101 GCTGTGCTCAGAGAGGCGACTGCACTTGTGTGTGTCACAGGCGCAACAGTCTCGAGAGG 2160  
DB 2101 GCTGTGCTCAGAGAGGCGACTGCACTTGTGTGTGTCACAGGCGCAACAGTCTCGAGAGG 2160  
QY 2161 CATCTTCTCTCTCTGTGAATCAGACAGCGGAGACCCAGCGCCCAAAATTTTCAAG 2220  
DB 2161 CATCTTCTCTCTCTGTGAATCAGACAGCGGAGACCCAGCGCCCAAAATTTTCAAG 2220  
QY 2221 GCAGAGTTTAAAGATGT 2280  
DB 2221 GCAGAGTTTAAAGATGT 2280  
QY 2281 TGTG 2284  
DB 2281 TGTG 2284

RESULT 4  
us-10-003-152-11  
; Sequence 11, Application US/10003152  
; Publication No. US20020151494A1  
; GENERAL INFORMATION:  
; APPLICANT: Shimkels, Richard  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Vernet, Corline  
; APPLICANT: Yang, Meijia  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Herrmann, John  
; TITLE OF INVENTION: No. US20020151494A1 Amino Acid Sequences for Human Semaphorin  
; FILE REFERENCE: 15966-554 Cura-54 CON-S12  
; CURRENT APPLICATION NUMBER: US/10/003,152  
; CURRENT FILING DATE: 2001-11-02  
; PRIOR APPLICATION NUMBER: 09/604,286  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 60/140,584  
; PRIOR FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2155  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (166)..(1935)  
US-10-003-152-11

Query Match 82.6%; Score 1887.4; DB 13; Length 2155;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 1 CGGCGCTTTCACACTCTGCTGATGTGAGAGCGGGTTTGGGTTCTGCAAGGCT 60  
DB 1 CGGCGCTTTCACACTCTGCTGATGTGAGAGCGGGTTTGGGTTCTGCAAGGCT 60  
QY 61 ATTGTCTGCGCTGGGGAAGGAGACAGGCGGAGACCGGGAGCTCGCTCGAGCGGCGC 120  
DB 61 ATTGTCTGCGCTGGGGAAGGAGACAGGCGGAGACCGGGAGCTCGCTCGAGCGGCGC 120

OY	121	ACGACGAGACACGCTGGGCTGAACTCAGAGGCGGGGGTGGCGCATGGCCCCACACACG	180
Db	121	ACGACGAGACAGCTGGGCTGAACTCAGAGGCGGGGGTGGCGCATGGCCCCACACACG	180
OY	181	GCTGTCTGGCTGTGGCAGCAAGGCTGTGGGGCTTGGGCAATTTGGGCTGTGGTGG	240
Db	181	GCTGTCTGGCTGTGGCAGCAAGGCTGTGGGGCTTGGGCAATTTGGGCTGTGGTGG	240
OY	241	AACCTTGTGGCGGCTTAAGACAGTGTCTTGTGGGAGCTGTGGCCACGGTAGTACGGGGGTTT	300
Db	241	AACCTTGTGGCGGCTTAAGACAGTGTCTTGTGGGAGCTGTGGCCACGGTAGTACGGGGGTTT	300
OY	301	TCCGACAGCCGGCATTCAGAGACTTCTGTACACTGACGAGCTGACGAGCCCACTGGGCTTCTG	360
Db	301	TCCGACAGCCGGCATTCAGAGACTTCTGTACACTGACGAGCTGACGAGCCCACTGGGCTTCTG	360
OY	361	TACGTGGCGCCCGCAGAGGCGCCCTGTTTGCTTCAGCATGAGAGCGCCCTGGAGCTGCAGGA	420
Db	361	TACGTGGCGCCCGCAGAGGCGCCCTGTTTGCTTCAGCATGAGAGCGCCCTGGAGCTGCAGGA	420
OY	421	GGGATCTCTGGGAGAGCCCGGTGGAGAAAGACTGTGATGTATCCAGAAAGGAAAGAAC	480
Db	421	GGGATCTCTGGGAGAGCCCGGTGGAGAAAGACTGTGATGTATCCAGAAAGGAAAGAAC	480
OY	481	AACCGACCGAGTGTCTTAACCTTCACTCGGCTCTTCGACAGCCCTACAAATGCTCCCACTG	540
Db	481	AACCGACCGAGTGTCTTAACCTTCACTCGGCTCTTCGACAGCCCTACAAATGCTCCCACTG	540
OY	541	TACGCTGTGGCACCCTAGCGCTTCCAGGCCAAGGTACACTACGTAAACATGTGTACACCTTC	600
Db	541	TACGCTGTGGCACCCTAGCGCTTCCAGGCCAAGGTACACTACGTAAACATGTGTACACCTTC	600
OY	601	ACTTTGGAGCATGAGAGTTTGAAGATGGGAAGGGCAAGTGTCCCTATGACCACTAAG	660
Db	601	ACTTTGGAGCATGAGAGTTTGAAGATGGGAAGGGCAAGTGTCCCTATGACCACTAAG	660
OY	661	GGCCATGTGGCTTCTTGTGGATGGTAGCTGTACTCGGCCACACTCAACACTTCTGT	720
Db	661	GGCCATGTGGCTTCTTGTGGATGGTAGCTGTACTCGGCCACACTCAACACTTCTGT	720
OY	721	GGCAGGAGAACCAATATCTGGCTTAACATGGGGGGCCCGACCACTGCATGAAAGACAGATAC	780
Db	721	GGCAGGAGAACCAATATATCTGGCTTAACATGGGGGGCCCGACCACTGCATGAAAGACAGATAC	780
OY	781	CTGGCGTTTGGCTTCAAGCAACTCTGATGAGGCTGTGCCTATGATACCTGAGAGTGTG	840
Db	781	CTGGCGTTTGGCTTCAAGCAACTCTGATGAGGCTGTGCCTATGATACCTGAGAGTGTG	840
OY	841	GGCAGCTTCAGGGGGAGCAGACAAAGTCTACTTCTTCTTAGGGAGCGGGCACTGGAG	900
Db	841	GGCAGCTTCAGGGGGAGCAGACAAAGTCTACTTCTTCTTAGGGAGCGGGCACTGGAG	900
OY	901	TCCGACTGTATGCCGACAGGTGTGTGCTCGTGTGGCCCGTGTGCAAGGGGCATATAG	960
Db	901	TCCGACTGTATGCCGACAGGTGTGTGCTCGTGTGGCCCGTGTGCAAGGGGCATATAG	960
OY	961	GGGGGCGCAGGAGCCCTGCAGAGAGAGTGTGACACGTTCTTGAAGGGCGGCTGGCATGC	1020
Db	961	GGGGGCGCAGGAGCCCTGCAGAGAGAGTGTGACACGTTCTTGAAGGGCGGCTGGCATGC	1020
OY	1021	TCTGCGCCCGAAGTGGCAGCTCTACTTCAACACGCTGCAAGGGCATCACACCTTGCAGGAC	1080
Db	1021	TCTGCGCCCGAAGTGGCAGCTCTACTTCAACACGCTGCAAGGGCATCACACCTTGCAGGAC	1080
OY	1081	ACCTCTTGGCACAACACACACCTTCTTTGGGGTTTTTCAAGCAGCAGTGGGTGACATGTAC	1140
Db	1081	ACCTCTTGGCACAACACACACCTTCTTTGGGGTTTTTCAAGCAGCAGTGGGTGACATGTAC	1140
OY	1141	CTGTGTGGGCATCTGTGAGTACCAGTGTGGAAGAGATCCAGCGGGGTGTTGAGGGGCCCTAT	1200
Db	1141	CTGTGTGGGCATCTGTGAGTACCAGTGTGGAAGAGATCCAGCGGGGTGTTGAGGGGCCCTAT	1200

OY	1201	AAGAGCTACCAATAGAGAGAGCCAGAAAGTGGGAGCCCTCACTAGGACCTGTACCCAGCCCT	1260
Db	1201	AAGAGGTACCATGAGAGAAAGCCAGAAAGTGGGAGCCCTCACTAGGACCTGTACCCAGCCCT	1260
OY	1261	CGGCGTGGCTGGTGCATTTAACAACTGGCATCGGCGCCAGCCGGCTACACCAAGCTCCCTGGAG	1320
Db	1261	CGGCGTGGCTGGTGCATTTAACAACTGGCATCGGCGCCAGCCGGCTACACCAAGCTCCCTGGAG	1320
OY	1331	CTACCCGACAAACATCTCTCAACTTGTCTCAAGAAAGCAACCCGGTGTATGGAGAGCAGGTGGGG	1380
Db	1331	CTACCCGACAAACATCTCTCAACTTGTCTCAAGAAAGCAACCCGGTGTATGGAGAGCAGGTGGGG	1380
OY	1381	CCTGGGTGAGAGCCGCCCTGGTGTGTGAAGAAAGGACACCACTTGACACCCAGCTGGTGGCC	1440
Db	1381	CCTGGGTGAGAGCCGCCCTGGTGTGTGAAGAAAGGACACCACTTGACACCCAGCTGGTGGCC	1440
OY	1441	GACCGGGTTACAGAGACTTGATGGAGCCACTATACAGTGTCTTCATTATGGACAGAGAGAC	1500
Db	1441	GACCGGGTTACAGAGACTTGATGGAGCCACTATACAGTGTCTTCATTATGGACAGAGAGAC	1500
OY	1501	GAGCTGGCTCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTGTATTTGAGAGAGCTGCAG	1560
Db	1501	GAGCTGGCTCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTGTATTTGAGAGAGCTGCAG	1560
OY	1561	CTGTTTGACACAGGAGCCCATGAGAAAGCCTGGGTGCTATCTCAGAGCAAGAAGCTGGTCTTT	1620
Db	1561	CTGTTTGACACAGGAGCCCATGAGAAAGCCTGGGTGCTATCTCAGAGCAAGAAGCTGGTCTTT	1620
OY	1621	GCCGGCTCCGCTCTCAGACTGATGAGAGCTGCCCCGATGGCCGACCTGCATTAAGTATGCTCC	1680
Db	1621	GCCGGCTCCGCTCTCAGACTGATGAGAGCTGCCCCGATGGCCGACCTGCATTAAGTATGCTCC	1680
OY	1681	TGTGAGAGACTGTGTCTCTGCGCCGGGAGCCCTATTGCGCTGGAGGTCAACACACGAGCCG	1740
Db	1681	TGTGAGAGACTGTGTCTCTGCGCCGGGAGCCCTATTGCGCTGGAGGTCAACACACGAGCCG	1740
OY	1741	TGTGTGGCGGTGGGGGGGACCTTGAGATCTTACTGATCCAGAGCATGATGACCTCGGAG	1800
Db	1741	TGTGTGGCGGTGGGGGGGACCTTGAGATCTTACTGATCCAGAGCATGATGACCTCGGAG	1800
OY	1801	ACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAATACAGTACAGGCCCACTGCCAAAAAC	1860
Db	1801	ACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAAGTACAGGCCCACT--CCCAAAAAAC	1857
OY	1861	ATCACGGTGTGGGGGGGACACAGACCTGTGTGCTGCCCTGCACACTTCTCCAACTTTGGCC	1920
Db	1858	ATCACGGTGTGGGGGGGACACAGACCTGTGTGCTGCCCTGCACACTTCTCCAACTTTGGCC	1917
OY	1921	CTG 1923	
Db	1918	CGG 1920	

```

RESULT 5
US-10-002-050-11
: Sequence 11, Application US/10002050
: Publication No. US20030032095A1
: GENERAL INFORMATION:
: APPLICANT: Shimizu, Richard
: APPLICANT: Fernandes, Elma
: APPLICANT: Vernet, Corine
: APPLICANT: Yang, Meljia
: APPLICANT: Boldog, Ferenc
: APPLICANT: Herrmann, John
: TITLE OF INVENTION: No. US20030032095A1
: FILE REFERENCE: 15966-554 Cura-54 CON-S14
: CURRENT APPLICATION NUMBER: US/10/002,050
: CURRENT FILING DATE: 2001-11-02
: PRIOR APPLICATION NUMBER: 09/604,286
: PRIOR FILING DATE: 2000-06-22
: PRIOR APPLICATION NUMBER: 60/140,584
: PRIOR FILING DATE: 1999-06-23
: NUMBER OF SEQ ID NOS: 49

```

SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 2155  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (166)..(1935)  
US-10-002-050-11

Query Match 82.6% Score 1887.4; DB 14; Length 2155;  
Best Local Similarity 99.5% Pred. No. 0;  
Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

1 CGGCCCCCTTCACACTCCTGCTGCTGATGTGGAACGGGGTTGGGGTTCTGACAGGCT 60  
1 CGGCCCCCTTCACACTCCTGCTGCTGATGTGGAACGGGGTTGGGGTTCTGACAGGCT 60  
61 ATTGTCTGCGCTGGGGAAAGGGAGACAGCCGGGACCGGGACCTCGCTCGACGGCGCGC 120  
61 ATTGTCTGCGCTGGGGAAAGGGAGACAGCCGGGACCGGGACCTCGCTCGACGGCGCGC 120  
121 ACCGAGGAGACAGTGGGCTGGAAGCTGAGAGCCGGGGCGTGGCGCATGGCCCCACACTGG 180  
121 ACCGAGGAGACAGTGGGCTGGAAGCTGAGAGCCGGGGCGTGGCGCATGGCCCCACACTGG 180  
181 GCTGTCTGCGTGTGCGCAGCAAGGCTGTGGGGCTGTGGGCAATTGGGCGTGAAGTGTGTGG 240  
181 GCTGTCTGCGTGTGCGCAGCAAGGCTGTGGGGCTGTGGGCAATTGGGCGTGAAGTGTGTGG 240  
241 AACCTTGTGCGCGCTGAAGACAGTGTCTTCTGGGAGCTGGGCAAGGTAGTACGGCGGTTG 300  
241 AACCTTGTGCGCGCTGAAGACAGTGTCTTCTGGGAGCTGGGCAAGGTAGTACGGCGGTTG 300  
301 TCCGAGACCGGACATCCAGGAGCTTCTGACACTGACGCGGAGCCGACCTGGGCTTCTG 360  
301 TCCGAGACCGGACATCCAGGAGCTTCTGACACTGACGCGGAGCCGACCTGGGCTTCTG 360  
361 TACGTGGGCGCCGAGAGGCGCTTGGCTTTCAGTGAAGAGCCCTGAGAGCTGCAAGA 420  
361 TACGTGGGCGCCGAGAGGCGCTTGGCTTTCAGTGAAGAGCCCTGAGAGCTGCAAGA 420  
421 GCGATCTCTGCGGAGGCGCCCGCTGGAGAAAGACTGATGATCCAGAAAGGAGAAC 480  
421 GCGATCTCTGCGGAGGCGCCCGCTGGAGAAAGACTGATGATCCAGAAAGGAGAAC 480  
481 AACGAGACGAGTCTCAACTTCATCGGCTTCTGAGAGCCCTGCAANTGCCCTCCACCTG 540  
481 AACGAGACGAGTCTCAACTTCATCGGCTTCTGAGAGCCCTGCAANTGCCCTCCACCTG 540  
541 TACGTCGTGGCACTACGCGCTTCCAGCCCAAGTGAACCTGACATCAATGCTCTACCTTC 600  
541 TACGTCGTGGCACTACGCGCTTCCAGCCCAAGTGAACCTGACATCAATGCTCTACCTTC 600  
601 ACTTTGAGAGATGAGAGGTTTGAAGATGGAGAGGCAAGTGTCCCTATGACCCAGCTAAG 660  
601 ACTTTGAGAGATGAGAGGTTTGAAGATGGAGAGGCAAGTGTCCCTATGACCCAGCTAAG 660  
661 GGGCATGCTGGCGCTTCTTGGAGTGTGAGCTGTACTGCGGCGACAGCTCAACACTTCTCTG 720  
661 GGGCATGCTGGCGCTTCTTGGAGTGTGAGCTGTACTGCGGCGACAGCTCAACACTTCTCTG 720  
721 GGCAGGAGACCATATATCTGCTGCTAATGATGGGGCCCACTCACTCCATGAGAGAGATAC 780  
721 GGCAGGAGACCATATATCTGCTGCTAATGATGGGGCCCACTCACTCCATGAGAGAGATAC 780  
781 CTGGCCTTTTGGCTCAACAGAACTTCTTGAAGCTGTGCTATGTAAGTGAAGTGTG 840  
781 CTGGCCTTTTGGCTCAACAGAACTTCTTGAAGCTGTGCTATGTAAGTGAAGTGTG 840  
841 GGCAGCTTCAAGGGGGAGAGAGCAAGTCTACTTCTTCAAGGAGCGGGCAGTGTAG 900  
841 GGCAGCTTCAAGGGGGAGAGAGCAAGTCTACTTCTTCAAGGAGCGGGCAGTGTAG 900

901 TCCGACTGCTATGCCAGACAGGTGATGCTGTGGCCCGTGTCTGCAAGGGCGATATG 960  
901 TCCGACTGCTATGCCAGACAGGTGATGCTGTGGCCCGTGTCTGCAAGGGCGATATG 960  
961 GGGGGCGACAGGACCTGTGACAGAGAGTGAACACAGTTCTTGAAGGGCGGCTGGATGC 1020  
961 GGGGGCGACAGGACCTGTGACAGAGAGTGAACACAGTTCTTGAAGGGCGGCTGGATGC 1020  
1021 TCTGCCCCGAACTGGAGCTTACTTCAACAGCTGCAAGGCGATGACACCTCTGACAGAC 1080  
1021 TCTGCCCCGAACTGGAGCTTACTTCAACAGCTGCAAGGCGATGACACCTCTGACAGAC 1080  
1081 ACCTCTGGACACACACACCTTCTTGGGGTTTTCAGAGCAGTGGGGTGAATGTATG 1140  
1081 ACCTCTGGACACACACACCTTCTTGGGGTTTTCAGAGCAGTGGGGTGAATGTATG 1140  
1141 CTGTGCGGCATCTGTGATGATGATGAGATGACAGCGGATGTTGAGGGCCCTAT 1200  
1141 CTGTGCGGCATCTGTGATGATGATGAGATGACAGCGGATGTTGAGGGCCCTAT 1200  
1201 AAGGAGTACATGAGAGAGAGCCAGAGGAGCCGCTACACTGACCTGTACCCAGCCCT 1260  
1201 AAGGAGTACATGAGAGAGAGCCAGAGGAGCCGCTACACTGACCTGTACCCAGCCCT 1260  
1261 CGGCTGGCTGTGCTATTAACACTGGCATGGGCGCACAGGCTACACAGCTCCCTGAG 1320  
1261 CGGCTGGCTGTGCTATTAACACTGGCATGGGCGCACAGGCTACACAGCTCCCTGAG 1320  
1321 CTACCCGACAACTCTCACTTCTGTCAGAGAGACCCGCTGATGAGAGAGAGTGGGG 1380  
1321 CTACCCGACAACTCTCACTTCTGTCAGAGAGACCCGCTGATGAGAGAGAGTGGGG 1380  
1381 CTTCCCGACAACTCTCACTTCTGTCAGAGAGACCCGCTGATGAGAGAGAGTGGGG 1380  
1381 CTTCCCGACAACTCTCACTTCTGTCAGAGAGACCCGCTGATGAGAGAGAGTGGGG 1380  
1441 GACCGGGTTACAGAGACTTATGATGAGAGCCACCTATACAGTGTCTTCAATTTGAGAGAG 1500  
1441 GACCGGGTTACAGAGACTTATGATGAGAGCCACCTATACAGTGTCTTCAATTTGAGAGAG 1500  
1501 GCGTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTGATGAGAGAGCTGCAG 1560  
1501 GCGTGGCTGCTCAAGGCTGTGAGCCTGGGGCCCTGGGTTCACTGATGAGAGAGCTGCAG 1560  
1561 CTGTTTGACAGAGGCCCATGAGAGGCTGTGATCTGAGAGCAAGAACCTGCTCTT 1620  
1561 CTGTTTGACAGAGGCCCATGAGAGGCTGTGATCTGAGAGCAAGAACCTGCTCTT 1620  
1621 GCGGAGCTCCGCTCAAGCTGTGAGCTGCGCCGCTGGCGGAGCTCATTAATATATGCTCC 1680  
1621 GCGGAGCTCCGCTCAAGCTGTGAGCTGCGCCGCTGGCGGAGCTCATTAATATATGCTCC 1680  
1681 TGTGACAGACTGTGCTGCGCCGAGACCCCTATTTGCGCTGAGAGCTCAACAGAGCCGC 1740  
1681 TGTGACAGACTGTGCTGCGCCGAGACCCCTATTTGCGCTGAGAGCTCAACAGAGCCGC 1740  
1741 TGTGAGCGCGTGGGGGCGCACTTGGATCTTACGATCCAGAGTGTGATGATGATGATGATG 1800  
1741 TGTGAGCGCGTGGGGGCGCACTTGGATCTTACGATCCAGAGTGTGATGATGATGATGATG 1800  
1801 ACTTCAGGCAATTTGCAACCTCGTGGAGTGAATACAGTGAAGGCCCACTTCCCAAAAC 1860  
1801 ACTTCAGGCAATTTGCAACCTCGTGGAGTGAATACAGTGAAGGCCCACTTCCCAAAAC 1860  
1861 ATCAGGGTGTGGGGGAGACAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
1861 ATCAGGGTGTGGGGGAGACAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
1921 CTG 1923  
1921 CTG 1923  
1918 CGG-1920

```
RESULT 6
US-10-002-304-11
/ Sequence 11, Application US/10002304
/ Publication No. US20030036185A1
/ GENERAL INFORMATION:
/ APPLICANT: Shlens, Richard
/ APPLICANT: Fernandes, Elma
/ APPLICANT: Vernet, Corine
/ APPLICANT: Yarn, Melja
/ APPLICANT: Boldoy, Ferenc
/ APPLICANT: Herrmann, John
/ TITLE OF INVENTION: Polynucleotides and polypeptides encoded thereby
/ FILE REFERENCE: 15366-54 Cura-54 CON-58
/ CURRENT APPLICATION NUMBER: US/10/002,304
/ PRIOR FILING DATE: 2001-11-02
/ PRIOR APPLICATION NUMBER: 09/604,286
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 60/140,584
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: Patent Ver. 2.0
/ SEQ ID NO 11
/ LENGTH: 2155
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (166)..(1935)
US-10-002-304-11

Query Match      82.6%; Score 1887.4; DB 14; Length 2155;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1914; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY      1  GGGCCCTCTCAGACCTCGCCCTGCTGATGGAAGGGGGTTGGGGTTGCGAGGCT 60
DB      1  GGGCCCTCTCAGACCTCGCCCTGCTGATGGAAGGGGGTTGGGGTTGCGAGGCT 60
QY      61  ATTGTCGGCTGGGGAAGGAGGAGCGCGGAGCTCCGCTCGACACCGCGCC 120
DB      61  ATTGTCGGCTGGGGAAGGAGGAGCGCGGAGCTCCGCTCGACACCGCGCC 120
QY      62  ACCAGCAGACAGCTGGCTGAACTCAGAGCCGGGGCTGCGCATGCGCCACACTGG 180
DB      62  ACCAGCAGACAGCTGGCTGAACTCAGAGCCGGGGCTGCGCATGCGCCACACTGG 180
QY      121 ACCAGCAGACAGCTGGCTGAACTCAGAGCCGGGGCTGCGCATGCGCCACACTGG 180
DB      121 ACCAGCAGACAGCTGGCTGAACTCAGAGCCGGGGCTGCGCATGCGCCACACTGG 180
QY      181 GCTGTCGCTGCTGGCAGCAGAGGCTGTGGGGCTGGGGCATTTGGGGCTGAGGTGGTGG 240
DB      181 GCTGTCGCTGCTGGCAGCAGAGGCTGTGGGGCTGGGGCATTTGGGGCTGAGGTGGTGG 240
QY      241 AACCTTGTCCCGCTAAGACAGTGTCTTGGGAGCTGGCCACGCTAGTACGGGCTTC 300
DB      241 AACCTTGTCCCGCTAAGACAGTGTCTTGGGAGCTGGCCACGCTAGTACGGGCTTC 300
QY      241 AACCTTGTCCCGCTAAGACAGTGTCTTGGGAGCTGGCCACGCTAGTACGGGCTTC 300
DB      241 AACCTTGTCCCGCTAAGACAGTGTCTTGGGAGCTGGCCACGCTAGTACGGGCTTC 300
QY      301 TCCAGACAGCGGATCCAGAGACTTCTGACACTGACGCTGACGAGCCCACTGGCTTCG 360
DB      301 TCCAGACAGCGGATCCAGAGACTTCTGACACTGACGCTGACGAGCCCACTGGCTTCG 360
QY      301 TCCAGACAGCGGATCCAGAGACTTCTGACACTGACGCTGACGAGCCCACTGGCTTCG 360
DB      301 TCCAGACAGCGGATCCAGAGACTTCTGACACTGACGCTGACGAGCCCACTGGCTTCG 360
QY      361 TACGTGGGGGCGCGAGAGGCGCTGTGGCTTACAGATGAGAGGCGCTGAGTGAAGA 420
DB      361 TACGTGGGGGCGCGAGAGGCGCTGTGGCTTACAGATGAGAGGCGCTGAGTGAAGA 420
QY      421 GCGATCTCTCGGAGAGCGCCCGCTGGAGAGAGAGACTGATGATATCCAGAAAGGAAAG 480
DB      421 GCGATCTCTCGGAGAGCGCCCGCTGGAGAGAGAGACTGATGATATCCAGAAAGGAAAG 480
QY      481 AACCAAGACGAGTGTCTTCACTTCACTCGCTCTGACGCTTACAAATGCTCCACCTG 540
DB      481 AACCAAGACGAGTGTCTTCACTTCACTCGCTCTGACGCTTACAAATGCTCCACCTG 540
QY      541 TACGCTGTGGGACTAGCGCTTCCAGCCCAAGTGCAGCTAGTCAACATGCTCACCTTC 600
DB      541 TACGCTGTGGGACTAGCGCTTCCAGCCCAAGTGCAGCTAGTCAACATGCTCACCTTC 600
```

```
DB      541 TACGCTGTGGGACTAGCGCTTCCAGCCCAAGTGCAGCTAGTCAACATGCTCACCTTC 600
QY      601 ACTTGGAGCATGAGAGATTGAGATGAGAGGAGCAAGTGTCCCTATGACCCACTAAG 660
DB      601 ACTTGGAGCATGAGAGATTGAGATGAGAGGAGCAAGTGTCCCTATGACCCACTAAG 660
QY      661 GGCATGCTGCGCTTCTGTGGATGATGAGTGTACTGCGGCACACTCAACACTTCCTG 720
DB      661 GGCATGCTGCGCTTCTGTGGATGATGAGTGTACTGCGGCACACTCAACACTTCCTG 720
QY      721 GGCAGGAAACCATTAATCTGCTTAACATGGGGGCGCCACACTCATGAAGCAAGATAG 780
DB      721 GGCAGGAAACCATTAATCTGCTTAACATGGGGGCGCCACACTCATGAAGCAAGATAG 780
QY      781 CTGGCTTTTGGCTCAAGCAACCTCACTTGTAGGCTGCTGCTATGATGACCTGAGATGG 840
DB      781 CTGGCTTTTGGCTCAAGCAACCTCACTTGTAGGCTGCTGCTATGATGACCTGAGATGG 840
QY      841 GGCAGCTTACGGGGGAGCAGCAGCAAGTCTACTTCTTCTTCAGGAGCGGCACTGGAG 900
DB      841 GGCAGCTTACGGGGGAGCAGCAGCAAGTCTACTTCTTCTTCAGGAGCGGCACTGGAG 900
QY      901 TCCGACTGCTATGCGGAGAGGTGGTGGCTGTGGCGCCGCTGCTGCAAGGCGCATATG 960
DB      901 TCCGACTGCTATGCGGAGAGGTGGTGGCTGTGGCGCCGCTGCTGCAAGGCGCATATG 960
QY      961 GGGGCGCAGCAGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB      961 GGGGCGCAGCAGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY      1021 TCTGCCCCGAACTGGCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080
DB      1021 TCTGCCCCGAACTGGCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080
QY      1081 ACCCTGGCAGCAACACACCTTCTTGGGGTTTCAAGCAGCAGTGGGGTACATGTAC 1140
DB      1081 ACCCTGGCAGCAACACACCTTCTTGGGGTTTCAAGCAGCAGTGGGGTACATGTAC 1140
QY      1141 CTGTCGGCCATCTGTAGTACAGAGTGGAGAGATCCAGCGGGGTGGAGGCCCTAT 1200
DB      1141 CTGTCGGCCATCTGTAGTACAGAGTGGAGAGATCCAGCGGGGTGGAGGCCCTAT 1200
QY      1201 AAGAGATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB      1201 AAGAGATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY      1261 GCGCTGGCTGTGATTAACAACTGGCAGCGCAGCGCTACACCAAGCTCCCTGGAG 1320
DB      1261 GCGCTGGCTGTGATTAACAACTGGCAGCGCAGCGCTACACCAAGCTCCCTGGAG 1320
QY      1321 CTACCCGACACATCTCAACTTCTGCAAGAGACACCGCTGATGAGAGAGAGAGAGAG 1380
DB      1321 CTACCCGACACATCTCAACTTCTGCAAGAGAGACACCGCTGATGAGAGAGAGAGAG 1380
QY      1381 CCTGGTGGAGCGCGCCCTGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB      1381 CCTGGTGGAGCGCGCCCTGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY      1441 GACCGGTTACAGAGCTTATGATGAGAGACCTTATACAGTGTCTTCAATGGGACAGAGAG 1500
DB      1441 GACCGGTTACAGAGCTTATGATGAGAGACCTTATACAGTGTCTTCAATGGGACAGAGAG 1500
QY      1501 GCGTGGCTGCTCAAGGCTGTGAGCGTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB      1501 GCGTGGCTGCTCAAGGCTGTGAGCGTGGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY      1561 CTGTTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB      1561 CTGTTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY      1621 GCGGCTGCGCGCTGCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB      1621 GCGGCTGCGCGCTGCTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
```



```

Oy 1321 CTACCCGACACATCTCACTTCTGTCAGAGAACCCCGCTGATGAGAGACAGTGGGG 1380
Db 1321 CTACCCGACACATCTCACTTCTGTCAGAGAACCCCGCTGATGAGAGACAGTGGGG 1380
Oy 1381 CCTGGGTGAGACCCCGCCCTCTCTGTGAAGAAAGGACCAACTTACCCACCTGGTGGC 1440
Db 1381 CCTGGGTGAGACCCCGCCCTCTCTGTGAAGAAAGGACCAACTTACCCACCTGGTGGC 1440
Oy 1441 GACCGGTTACAGAGACTGATGAGAGCCATATACAGTGTCTTATTCATTGGACAGAGAGAC 1500
Db 1441 GACCGGTTACAGAGACTGATGAGAGCCATATACAGTGTCTTATTCATTGGACAGAGAGAC 1500
Oy 1501 GCGTGGCTGTCAAGAGCTGTGAGCCCTGGGCGCTGGGTTACAGTATGAGAGCTGACAG 1560
Db 1501 GCGTGGCTGTCAAGAGCTGTGAGCCCTGGGCGCTGGGTTACAGTATGAGAGCTGACAG 1560
Oy 1561 CTGTTTACAGAGAGCCCATATAGAAAGCTGTGCTATCTCAGAGACCAAGACCTGCTTT 1620
Db 1561 CTGTTTACAGAGAGCCCATATAGAAAGCTGTGCTATCTCAGAGACCAAGACCTGCTTT 1620
Oy 1621 GCCGGCTCCCGCTCTCAGAGCTGTGAGCTGCGCTGGGCGAGCTGATTAAGTATCGCTCC 1680
Db 1621 GCCGGCTCCCGCTCTCAGAGCTGTGAGCTGCGCTGGGCGAGCTGATTAAGTATCGCTCC 1680
Oy 1681 TGTGACAGACTGTGTCTGCGCCGGGAGCCCTATTGCGGCTGAGAGCTGACACAGAGCCGC 1740
Db 1681 TGTGACAGACTGTGTCTGCGCCGGGAGCCCTATTGCGGCTGAGAGCTGACACAGAGCCGC 1740
Oy 1741 TGTGAGCGCGTGGGTGGGCACTTTGGATCTTACTGATCCAGACATGATGATACCTCGGAC 1800
Db 1741 TGTGAGCGCGTGGGTGGGCACTTTGGATCTTACTGATCCAGACATGATGATACCTCGGAC 1800
Oy 1801 ACTTCAGGACATTTTCAACCTCTGTCAGAGTAAAGTATACAGAGCCCACTCCCAAAAC 1860
Db 1801 ACTTCAGGACATTTTCAACCTCTGTCAGAGTAAAGTATACAGAGCCCACTCCCAAAAC 1860
Oy 1861 ATCAGCGTGTGGGGGACAGAGACCTGTGCTGCGCCCTGACCTCTCTCCAACTTGGCC 1920
Db 1861 ATCAGCGTGTGGGGGACAGAGACCTGTGCTGCGCCCTGACCTCTCTCCAACTTGGCC 1920
Oy 1921 CTG 1923
Db 1918 CCG 1920

```

RESULT 8  
US-10-002-050-21  
Sequence 21, Application US/10002050  
Publication No. US20030032095A1

GENERAL INFORMATION:  
APPLICANT: Shinkels, Richard  
APPLICANT: Fernandez, Elma  
APPLICANT: Vernet, Corine  
APPLICANT: Yang, Melja  
APPLICANT: Boldog, Ferenc  
APPLICANT: Heitmann, John

TITLE OF INVENTION: No. US20030032095A1el Nucleic Acid Sequences Encoding Human Sema4  
FILE REFERENCE: 15966-554 Cura-54 CON-S14  
CURRENT APPLICATION NUMBER: US/10/002,050  
CURRENT FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: 09/604,286  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 60/140,584  
PRIOR FILING DATE: 1999-06-23  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 21  
LENGTH: 2156  
TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS

LOCATION: (166)..(2037)  
NAME/KEY: variation  
LOCATION: (1)..(2156)  
OTHER INFORMATION: N may be any nucleotide  
US-10-002-050-21

Query Match 82.3%; Score 1879.4; DB 14; Length 2156;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1909; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

```

Oy 1 CGGGCCCTTCACACCTCTCGCCCTGCTGATATGGAACGGGGTTGGGGTTCTGCAGGGCT 60
Db 1 CGGGCCCTTCACACCTCTCGCCCTGCTGATATGGAACGGGGTTGGGGTTCTGCAGGGCT 60
Oy 61 ATTGCTGCGCTGGGGAAGGAGAGAGCCGGAGACCTCCGCTGCAGAGCGGCGC 120
Db 61 ATTGCTGCGCTGGGGAAGGAGAGAGCCGGAGACCTCCGCTGCAGAGCGGCGC 120
Oy 121 ACCAGCAGAGAGCTGCGCTGAAGCTCAGAGCCGGGGCGTGGCCCATGGCCCACTGG 180
Db 121 ACCAGCAGAGAGCTGCGCTGAAGCTCAGAGCCGGGGCGTGGCCCATGGCCCACTGG 180
Oy 181 GCTGCTGGGCTGTGGGACAGAAAGCTGTGGGGCGCTGGGCAATTGGGCGTGAAGTGG 240
Db 181 GCTGCTGGGCTGTGGGACAGAAAGCTGTGGGGCGTGGGCAATTGGGCGTGAAGTGG 240
Oy 241 AACCTGTGCGCGTGAAGACAGTGTCTTCTGGGAGCTGGGCGAGGTAGTACGGCGTTC 300
Db 241 AACCTGTGCGCGTGAAGACAGTGTCTTCTGGGAGCTGGGCGAGGTAGTACGGCGTTC 300
Oy 301 TCCAGAGCCGCAATCCAGAGACTTCTGACACTGACAGCTGACAGAGCCCACTGGGCTTCG 360
Db 301 TCCAGAGCCGCAATCCAGAGACTTCTGACACTGACAGCTGACAGAGCCCACTGGGCTTCG 360
Oy 361 TACGTTGGGCGCCGAGAGAGCCCTGTTGCTTCAAGCATGAGAGGCGCTGGAGCTGAAGA 420
Db 361 TACGTTGGGCGCCGAGAGAGCCCTGTTGCTTCAAGCATGAGAGGCGCTGGAGCTGAAGA 420
Oy 421 GCGATCTCTCTGGGAGGCGCCCGTGGAGAGAAAGAGTGTATCCAGAAAGGAGAAC 480
Db 421 GCGATCTCTCTGGGAGGCGCCCGTGGAGAGAAAGAGTGTATCCAGAAAGGAGAAC 480
Oy 481 AACAGACCGAGTGTCTTCAACTTCACTCGCTTCTGACAGCCCTTACATGCTTCCACTG 540
Db 481 AACAGACCGAGTGTCTTCAACTTCACTCGCTTCTGACAGCCCTTACATGCTTCCACTG 540
Oy 541 TACGTTGGGCGCCGAGAGAGCCCTTCCAGAGCCCAAGTGTGACATGATGATGATGATG 600
Db 541 TACGTTGGGCGCCGAGAGAGCCCTTCCAGAGCCCAAGTGTGACATGATGATGATGATG 600
Oy 601 ACTTGGAGCATGAGAGATTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
Db 601 ACTTGGAGCATGAGAGATTGAAGATGGGAAGGCAAGTGTCCCTATGACCCAGCTAAG 660
Oy 661 GCGCATGCTGGCCTTCTTGTGATGTGATGCTGTATCTGCGCACACTAACAACCTTCTG 720
Db 661 GCGCATGCTGGCCTTCTTGTGATGTGATGCTGTATCTGCGCACACTAACAACCTTCTG 720
Oy 721 GGCAGGGAACCATTTATCTGCTGTAACATGAGGCGCCCACTCATATGAACAGAGTAC 780
Db 721 GGCAGGGAACCATTTATCTGCTGTAACATGAGGCGCCCACTCATATGAACAGAGTAC 780
Oy 781 CTGGCTTTTGGCTCAAGAACTCACTTGTAGGCTGTGCTATGATCTGAGAGTGTG 840
Db 781 CTGGCTTTTGGCTCAAGAACTCACTTGTAGGCTGTGCTATGATCTGAGAGTGTG 840
Oy 841 GGCAGCTTCAGGGGGAGAGAGAGTGTACTTCTTCTCAGAGAGCGGGAGTGTAG 900
Db 841 GGCAGCTTCAGGGGGAGAGAGAGTGTACTTCTTCTCAGAGAGCGGGAGTGTAG 900
Oy 901 TCCAGCTGTATGCGAGAGAGTGTGTGCTGTGCGGCTGTCTGCAAGGCGATATG 960
Db 901 TCCAGCTGTATGCGAGAGAGTGTGTGCTGTGCGGCTGTCTGCAAGGCGATATG 960

```

QY	961	GGGGGCGAGCGACCCCTGCGACAGAGGAAGTGGACACAGTTCCTCGAAGGGCGGGCTGGCAGTC	1020
Db	961	GGGGGCGAGCGACCCCTGCGACAGAGGAAGTGGACACAGTTCCTCGAAGGGCGGGCTGGCAGTC	1020
QY	1021	TCTGCCCCGAACCTGGCAGCTCTACTTTCGAACAGATCGAGGGGATGCAACCCCTCAGAGAC	1080
Db	1021	TCTGCCCCGAACCTGGCAGCTCTACTTTCGAACAGATCGAGGGGATGCAACCCCTCAGAGAC	1080
QY	1081	ACCTCCCTGGCACAACACCACCTCTTTTGGGGTTTTTCAAGCACAGTGGGGTGCATGTAC	1140
Db	1081	ACCTCCCTGGCACAACACCACCTCTTTTGGGGTTTTTCAAGCACAGTGGGGTGCATGTAC	1140
QY	1141	CTGTGCGGCATCTGTGAGTATCAGTGTGGAAAGANCCAGCGGGGTGTGAGGGGCCCTAT	1200
Db	1141	CTGTGCGGCATCTGTGAGTATCAGTGTGGAAAGANCCAGCGGGGTGTGAGGGGCCCTAT	1200
QY	1201	AAGGAGTACCATGAGAGAAAGCCCAAGTGGGACCGCTACACTGCACCCCTGTACCCAGCCCT	1260
Db	1201	AAGGAGTACCATGAGAGAAAGCCCAAGTGGGACCGCTACACTGCACCCCTGTACCCAGCCCT	1260
QY	1261	CGGCTTGGCTGTGTGATTTACACTGGCATCGGCGCCACGGCTACACCAAGTCCCTGTGAG	1320
Db	1261	CGGCTTGGCTGTGTGATTTACACTGGCATCGGCGCCACGGCTACACCAAGTCCCTGTGAG	1320
QY	1321	CTACCCGCAACATCTCTCAACTTTCGTCAAGAAAGACCCGGCTGATGAGGAGACAGTGGGG	1380
Db	1321	CTACCCGCAACATCTCTCAACTTTCGTCAAGAAAGACCCGGCTGATGAGGAGACAGTGGGG	1380
QY	1381	CCCTGGTGGAGCCGCCCTGCTGCTGTAAGAAAGGGCACCACTTCACCCACCTGGTGGCC	1440
Db	1381	CCCTGGTGGAGCCGCCCTGCTGCTGTAAGAAAGGGCACCACTTCACCCACCTGGTGGCC	1440
QY	1441	GACCGGGTTACAGAGACTTGATGAGAGCCACTATACAGTGTCTGTTCATTGGCACAGAGAC	1500
Db	1441	GACCGGGTTACAGAGACTTGATGAGAGCCACTATACAGTGTCTGTTCATTGGCACAGAGAC	1500
QY	1501	GGCTGGCTGCTCAAGGCGTGTGAGCCTGGGGCCCTGGGTTCACTGATTGAGAGACTGCAG	1560
Db	1501	GGCTGGCTGCTCAAGGCGTGTGAGCCTGGGGCCCTGGGTTCACTGATTGAGAGACTGCAG	1560
QY	1561	CTGTTTTCACGAGAGCCCATGAGAGACCTGGTGTATCTCAGAGCAAGAACCTGCTCTTT	1620
Db	1561	CTGTTTTCACGAGAGCCCATGAGAGACCTGGTGTATCTCAGAGCAAAACCTGCTCTTT	1620
QY	1621	GGCGGCTCCGCGCTCTCAGCTGGTGCAGAGTGGCCCTGGCGCATCTGATTAAGTATGCGTCC	1680
Db	1621	GGCGGCTCCGCGCTCTCAGCTGGTGCAGAGTGGCCCTGGCGCATCTGATTAAGTATGCGTCC	1680
QY	1681	TGTGCAGAGCTGTGTCCTGCGCCCGGGACCCCTATTTCGCGCTGGAGCGTCAACACGCGCG	1740
Db	1681	TGTGCAGAGCTGTGTCCTGCGCCCGGGACCCCTATTTCGCGCTGGAGCGTCAACACGCGCG	1740
QY	1741	TGTGTGGCCGTGGGGTGGCCACTTTTGGATCTTTACTGATCAGACATGTGATGACTCGGAC	1800
Db	1741	TGTGTGGCCGTGGGGTGGCCACTTTTGGATCTTTACTGATCAGACATGTGATGACTCGGAC	1800
QY	1801	ACTTCAGGCACTTTCGAACCTCGGTGGCAGTAAATACAGTAAAGGCCCATCTCCCAAAAAC	1860
Db	1801	ACTTCAGGCACTTTCGAACCTCGGTGGCAGTAAAGTAAAGGCCCATCT--CCCAAAAAC	1857
QY	1861	ATCACGGTGTGGCGGGCGACAGACCTGGTGTGCTCCGTGCAACCTCTCCCTCAACTTGGCC	1920
Db	1858	ATCACGGTGTGGCGGGCGACAGACCTGGTGTGCTCCGTCCACCTCTCTCCACTTGGCC	1917
QY	1921	CTG 1923	
Db	1918	CGG 1920	

```

? Publication No. US20030036185A1
? GENERAL INFORMATION:
?
? APPLICANT: Shimkets, Richard
? APPLICANT: Fernandes, Elma
? APPLICANT: Vernet, Corine
? APPLICANT: Yang, Melja
? APPLICANT: Boldog, Ferenc
? APPLICANT: Herrmann, John
?
? TITLE OF INVENTION: polynucleotides and polypeptides encoded thereby
?
? FILE REFERENCE: 15966-554 Cura-54 CON-58
?
? CURRENT FILING DATE: 2001-11-02
?
? PRIOR APPLICATION NUMBER: 09/604,286
?
? PRIOR FILING DATE: 2000-06-22
?
? PRIOR APPLICATION NUMBER: 60/7140,584
?
? PRIOR FILING DATE: 1999-06-23
?
? NUMBER OF SEQ ID NOS: 49
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 21
?
? LENGTH: 2156
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
? FEATURE:
?
? NAME/KEY: CDS
?
? LOCATION: (166)..(2037)
?
? NAME/KEY: variation
?
? LOCATION: (1)..(2156)
?
? OTHER INFORMATION: N may be any nucleotide
?
US-10-002-304-21

```

Query Match	82.3%	Score 1879.4	DB 14	Length 2156
Best Local Similarity	99.3%	Pred. No. 0		
Matches 1909	Conservative	0	Mismatches 11	Indels 3
				Gaps 2

RESULT 9  
US-10-002-304-21  
; Sequence 21, Application US/10002304

OY	1	CGGCGCTCTCACACTCCCTGGCCGCTATGTGGAACGGGGTTTGGGGTTCTGGAGGGCT	60
Db	1	CGGCGCTTCTACACTCCTGCGCTGCTATGTGGAACGGGGTTTGGGGTTCTGGAGGGCT	60
OY	61	ATTGTCTGCGCTGGGGGAAGGGGAGACAGCCGGGACCTCCGCTTCGACAGCCGCGC	120
Db	61	ATTGTCTGCGCTGGGGGAAGGGGAGACAGCCGGGACCTCCGCTTCGACAGCCGCGC	120
OY	121	ACGACGACGACAGCTGGCCTTAAGCTCAGAGCCGGGGGCTGCCCATGCCCCACACTGG	180
Db	121	ACGACGACGACAGCTGGCCTTAAGCTCAGAGCCGGGGGCTGCCCATGCCCCACACTGG	180
OY	181	GCTGCTGGCTGCTGGCAGCAAGGCTGGGGGCTGGGCAATTGGGGGCTGAGTGTGGTGG	240
Db	181	GCTGCTGGCTGCTGGCAGCAAGGCTGGGGGCTGGGCAATTGGGGGCTGAGTGTGGTGG	240
OY	241	AACCTTGTGCGCGCTAAGACAGTGTCTTCTGGGGAGCTGGCCACAGGTAGTACGGCGTTTC	300
Db	241	AACCTTGTGCGCGCTAAGACAGTGTCTTCTGGGGAGCTGGCCACAGGTAGTACGGCGTTTC	300
OY	301	TCCGACAGCCGGCATTCAGAGACTTCTGTACACTGACGCTGACGGAGGCCACTGGCCTTGTG	360
Db	301	TCCGACAGCCGGCATTCAGAGACTTCTGTACACTGACGCTGACGGAGGCCACTGGCCTTGTG	360
OY	361	TACGTTGGGCGCCCGAGAGGAGCCCTGTTTACCTTCAGACATGAGAGGCCCTGGACCTGCAAGA	420
Db	361	TACGTTGGGCGCCCGAGAGGAGCCCTGTTTTCCTTCACACATGAGAGGCCCTTGGACCTGCAAGA	420
OY	421	GCGATCTCTCTGGAGGAGCCCCGCTGGAGAGAAAGTGAAGTATTCAGAAAGGGAGAAC	480
Db	421	GCGATCTCTCTGGAGGAGCCCCGCTGGAGAGAAAGTGAAGTATTCAGAAAGGGAGAAC	480
OY	481	AACCAAGACCGAGTGTCTTAATTATCCGCTTCTGCAAGCCCTTCAATATGCTCCCACTGG	540
Db	481	AACCAAGACCGAGTGTCTTAATTATCCGCTTCTGCAAGCCCTTCAATATGCTCCCACTGG	540
OY	541	TACGTTGTGGACACTACGCTTCAGGCCCAAGAGACACTACGTCACAACTGCTCAACCTTC	600
Db	541	TACGTTGTGGACACTACGCTTCAGGCCCAAGAGACACTACGTCACAACTGCTCAACCTTC	600



```

QY 601 ACTTTGAGCATGAGAGCTTTGAAAGATGGGAAAGGTGTCCCTATGACCCAGCTAAG 660
DB 601 ACTTTGAGCATGAGAGATTTGAAAGATGGGAAAGGTGTCCCTATGACCCAGCTAAG 660
QY 661 GGGCATTGGTGGCCCTTCTGTGATGTTGAGTGTACTGTGGGACACACTCAAACTCTCTG 720
DB 661 GGGCATTGGTGGCCCTTCTGTGATGTTGAGTGTACTGTGGGACACACTCAAACTCTCTG 720
QY 721 GGCAGGAAACCATTTATCTGTGCTAATCATGAGGCCCCACACTCCATGAAAGACAGTAC 780
DB 721 GGCAGGAAACCATTTATCTGTGCTAATCATGAGGCCCCACACTCCATGAAAGACAGTAC 780
QY 781 CTGGCTTTTGGCTTAACGAACCTGACTTTGTAGCTCTGCCATATGACCTGAGAGTGTG 840
DB 781 CTGGCTTTTGGCTTAACGAACCTGACTTTGTAGCTCTGCCATATGACCTGAGAGTGTG 840
QY 841 GGCAGCTTCAGGGGGGAGAGACAAAGGTACTCTTCTTCAGGAGCGGGCACTGGAG 900
DB 841 GGCAGCTTCAGGGGGGAGAGACAAAGGTACTCTTCTTCAGGAGCGGGCACTGGAG 900
QY 901 TCCGACTCTATGCGGAGCAGGTGGTGGCTGTGGCCCGTGTGCAAGGGCGATATG 960
DB 901 TCCGACTCTATGCGGAGCAGGTGGTGGCTGTGGCCCGTGTGCAAGGGCGATATG 960
QY 961 GGGGGCGACGAGACCTTCAGAGAGAGTGGACACGTTCTGAGAGCGCGCTGGCATGC 1020
DB 961 GGGGGCGACGAGACCTTCAGAGAGAGTGGACACGTTCTGAGAGCGCGCTGGCATGC 1020
QY 1021 TCTGCCCGGAACTGAGCTGTACTTCAACGAGCTGAGGGGATGACACCCCTGAGAGAC 1080
DB 1021 TCTGCCCGGAACTGAGCTGTACTTCAACGAGCTGAGGGGATGACACCCCTGAGAGAC 1080
QY 1081 AACTCTGGGCAACACACACCTCTTTGGGGTTTTTCAAGCAGTGGGGTGCATGTAC 1140
DB 1081 AACTCTGGGCAACACACACCTCTTTGGGGTTTTTCAAGCAGTGGGGTGCATGTAC 1140
QY 1141 CTGTGGGCACTGTGTGATGACAGTGTGAAAGATGCCAGGGGTGTTGAGGGCCCTAT 1200
DB 1141 CTGTGGGCACTGTGTGATGACAGTGTGAAAGATGCCAGGGGTGTTGAGGGCCCTAT 1200
QY 1201 AAGGAGTACATGAGAGAGACCCAGAGTGGGACCGCTGACTGACCTGTACCCAGCCCT 1260
DB 1201 AAGGAGTACATGAGAGAGACCCAGAGTGGGACCGCTGACTGACCTGTACCCAGCCCT 1260
QY 1261 GGGCGTGGCTGTGATTAAACAAGTGGCATGGGCGCACGCGCTACACAGCTCCCTGGAG 1320
DB 1261 GGGCGTGGCTGTGATTAAACAAGTGGCATGGGCGCACGCGCTACACAGCTCCCTGGAG 1320
QY 1321 CTACCCGACAAACATCTCTCAACTTGTCAAGAAAGCACCCGCTGATGAGAGAGAGTGGG 1380
DB 1321 CTACCCGACAAACATCTCTCAACTTGTCAAGAAAGCACCCGCTGATGAGAGAGAGTGGG 1380
QY 1381 CTTGCGTGGAGCGCGCCCTGTCTGTAAGAAAGGACCAACCTTACCCACACTGGTGGCC 1440
DB 1381 CTTGCGTGGAGCGCGCCCTGTCTGTAAGAAAGGACCAACCTTACCCACACTGGTGGCC 1440
QY 1441 GACCGGGTTACAGACTTGTATGAGAGCCACTATACAGTGTGTTTATTTGGCAGAGAGAC 1500
DB 1441 GACCGGGTTACAGACTTGTATGAGAGCCACTATACAGTGTGTTTATTTGGCAGAGAGAC 1500
QY 1501 GGGCTGGCTGTCAAGGCTGTGAGCGCTGGGGCCCTGGGTTACCTGATTTGAGAGCTGAG 1560
DB 1501 GGGCTGGCTGTCAAGGCTGTGAGCGCTGGGGCCCTGGGTTACCTGATTTGAGAGCTGAG 1560
QY 1561 CTGTTTGAACGAGACCCCATGAGAGCCTGTGTCTATCTCAGAGAGAAAGTGTCTTTT 1620
DB 1561 CTGTTTGAACGAGACCCCATGAGAGCCTGTGTCTATCTCAGAGAGAAAGTGTCTTTT 1620
QY 1621 GCGCGCTCCGCTCTCAGCTGTGTGAGCTGCGCTGGCGAGACTGCATAAAGTATGCTCC 1680
DB 1621 GCGCGCTCCGCTCTCAGCTGTGTGAGCTGCGCTGGCGAGACTGCATAAAGTATGCTCC 1680

```

```

QY 1681 TGTGAGACTGTGTCTCTGCGCCGGGACCCCTATTTGCGCCCTGAGGCTCAACACAGCCGC 1740
DB 1681 TGTGAGACTGTGTCTCTGCGCCGGGACCCCTATTTGCGCCCTGAGGCTCAACACAGCCGC 1740
QY 1741 TGTGTGGCGTGGGGGCGCACTTGGATCTTACATGATCCAGATGATGACCTGGAG 1800
DB 1741 TGTGTGGCGTGGGGGCGCACTTGGATCTTACATGATCCAGATGATGACCTGGAG 1800
QY 1801 ACTTAGGCAATTTGCAACCTCCGTGGCAGTAAGATACAGTACAGGCCCACTNCCAAAAAC 1860
DB 1801 ACTTAGGCAATTTGCAACCTCCGTGGCAGTAAG--AAAGTACGGCCACT--CCAAAAAC 1857
QY 1861 ATCAGCGTGTGGGGGCGCAGACACTGTGTGCTGCTGCGCCACTCTCTCCACTTTGGCC 1920
DB 1861 ATCAGCGTGTGGGGGCGCAGACACTGTGTGCTGCTGCGCCACTCTCTCCACTTTGGCC 1917
QY 1921 CTG 1923
DB 1918 CCG 1920

```

```

RESULT 10
US-10-149-819-25
; Sequence 25, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Valda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandira
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dying Alina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janlice
; APPLICANT: BUREFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 3293
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030044913A1 1674661CB1
US-10-149-819-25

```

```

Query Match 77.1%; Score 1761; DB 14; Length 3293;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1811; Conservative 0; Mismatches 45; Indels 3; Gaps 2;
QY 84 CAGGCGGGAGCGGGAGACCTCGCTGCGAGCGCGCGGACACACAGAGACAGTGGCGCTGA 143
DB 52 CCGAAGCGGAGGGCGCGCGCGCGGAGACTGCACTCGCGCGCAATTAAGAGACAGTGGCGCTGA 111
QY 144 GCTCAGAGCGGGGCGGTGCGCCATGAGCCCACTAGGGCTGTGCTGGTCTGGCAGAG 203
DB 112 GCTCAGAGCGGGGCGGTGCGCCATGAGCCCACTAGGGCTGTGCTGGTCTGGCAGAG 171
QY 204 GCTGTGGGCGCTGGGCAATTTGGGCTGAGGTGTGTGGAACCTTGTGCCGTAAGACAGT 263
DB 172 GCTGTGGGCGCTGGGCAATTTGGGCTGAGGTGTGTGGAACCTTGTGTGCCGTAAGACAGT 231
QY 264 GTCTCTGGGAGCGGGGAGAGGTAGTACGGGGGTTCTCCAGACCGGCAATCCAGAGACTT 323
DB 232 GTCTCTGGGAGCGGGGAGAGGTAGTACGGGGGTTCTCCAGACCGGCAATCCAGAGACTT 291

```



QY 324 CTTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCGCT 383  
DB 292 CTTGACACTGACGCTGACGAGGCCACTGGGCTTCTGTACGTGGGCGCCGAGAGGCGCT 351  
QY 384 GTTGGCTTACGATGAGAGGCGCTGGAGCTGCAAGAGGCGATCTCCCTGGAGGCGCCGCT 443  
DB 352 GTTGGCTTACGATGAGAGGCGCTGGAGCTGCAAGAGGCGATCTCCCTGGAGGCGCCGCT 411  
QY 444 GGAAGAAGAACTGAGTGTATCCAGAAAGGAAAGAACCAACAGACGAGTGTCTCAACTT 503  
DB 412 GGAAGAAGAACTGAGTGTATCCAGAAAGGAAAGAACCAACAGACGAGTGTCTCAACTT 471  
QY 504 CATCCGCTTCTGACGCTTACAAATGCTCCCACTGTACGTGTGGCACTTACGCTT 563  
DB 472 CATCCGCTTCTGACGCTTACAAATGCTCCCACTGTACGTGTGGCACTTACGCTT 531  
QY 564 CCAAGCCCAAGTGAACCTGACAAAGTCAAGTCACTTCACTTGGAGATGAGAGTGTGA 623  
DB 532 CCAAGCCCAAGTGAACCTGACAAAGTCAAGTCACTTCACTTGGAGATGAGAGTGTGA 591  
QY 624 AGATGGGAAGGCAAGTGTCTTATGACAGGCTAAGGCGCAATGCTGGCTTCTGTGGA 683  
DB 592 AGATGGGAAGGCAAGTGTCTTATGACAGGCTAAGGCGCAATGCTGGCTTCTGTGGA 651  
QY 684 TGGTGAAGTGTCTGAGGCACTCAACAACTTCTGGGCAAGGAAACCAATTATCTGCG 743  
DB 652 TGGTGAAGTGTCTGAGGCACTCAACAACTTCTGGGCAAGGAAACCAATTATCTGCG 711  
QY 744 TAAATATGAGGCGCCCACTCACTCAAGAGATGAGTACCTGGCTTGGCTTCAACGAAC 803  
DB 712 TAAATATGAGGCGCCCACTCACTCAAGAGATGAGTACCTGGCTTGGCTTCAACGAAC 771  
QY 804 TCACTTGTAGGCTCTGCTATGATGAGTGTGGGAGCTTCAAGGAGGAGGAGCA 863  
DB 772 TCACTTGTAGGCTCTGCTATGATGAGTGTGGGAGCTTCAAGGAGGAGGAGCA 831  
QY 864 CAAAGTCTACTTCTTTCAGGAGGCGGAGTGAAGTCCGACTGATGCGGAGCAGGT 923  
DB 832 CAAAGTCTACTTCTTTCAGGAGGCGGAGTGAAGTCCGACTGATGCGGAGCAGGT 891  
QY 924 GGTGGGCTGTGGGCGGCTGCTGCAAGGCGGAGATATGGGCGGCGCAGGAGCCTGCAAG 983  
DB 892 GGTGGGCTGTGGGCGGCTGCTGCAAGGCGGAGATATGGGCGGCGCAGGAGCCTGCAAG 951  
QY 984 GAAGTGAAGCAAGTCTCTGAAGGCGGCTGCGATGCTGCGCGAAGTGGCAGCTCTA 1043  
DB 952 GAAGTGAAGCAAGTCTCTGAAGGCGGCTGCGATGCTGCGCGAAGTGGCAGCTCTA 1011  
QY 1044 CTTCAACAGCTGACGAGGAGTGAACACCTGACAGGACACTCTGCGCAACACACACTT 1103  
DB 1012 CTTCAACAGCTGACGAGGAGTGAACACCTGACAGGACACTCTGCGCAACACACACTT 1071  
QY 1104 CTTGGGCTTTCAGGAGGAGTGAAGTGAATGCTGCGGCAATCTGTGAGTGA 1163  
DB 1072 CTTGGGCTTTCAGGAGGAGTGAAGTGAATGCTGCGGCAATCTGTGAGTGA 1131  
QY 1164 GTTGAAGAGATCAAGCGGCTGTTGAGGCGGCTTATGAAGAGATCAATGAAGAGCA 1223  
DB 1132 GTTGAAGAGATCAAGCGGCTGTTGAGGCGGCTTATGAAGAGATCAATGAAGAGCA 1191  
QY 1224 GAAAGTGAAGCTGACACTGAGCTGATACCAAGGCTGAGGCTGCTGATTAACAA 1283  
DB 1192 GAAAGTGAAGCTGACACTGAGCTGATACCAAGGCTGAGGCTGCTGATTAACAA 1251  
QY 1284 CTGGCATGCGGCGCAGGCTGACACAGCTGCTGAGACTACCGAACAATCTCAACTT 1343  
DB 1252 CTGGCATGCGGCGCAGGCTGACACAGCTGCTGAGACTACCGAACAATCTCAACTT 1311  
QY 1344 CGTCAAGAACACACCGCTGATGAGAGAGCAAGTGGGCGCTGGGTGAGCGCGCGCGCT 1403  
DB 1312 CGTCAAGAACACACCGCTGATGAGAGAGCAAGTGGGCGCTGGGTGAGCGCGCGCT 1371

QY 1404 CTTGAAGAGGAGGCAACCACTTACCAACCTGCTGCGGCGGAGGCTTACAGACTGTATG 1463  
DB 1372 CTTGAAGAGGAGGCAACCACTTACCAACCTGCTGCGGCGGAGGCTTACAGACTGTATG 1431  
QY 1464 AGCCACTATACATGCTGTTCATGTCACAGAGAGAGAGGCTGCTGCTCAAGCTGTGAG 1523  
DB 1432 AGCCACTATACATGCTGTTCATGTCACAGAGAGAGAGGCTGCTGCTCAAGCTGTGAG 1491  
QY 1524 CTTGGGCGGCTGAGTTCACCTGATTTGAGAGAGCTGACACTGTTTGAACAGAGGCGCATGAG 1583  
DB 1492 CTTGGGCGGCTGAGTTCACCTGATTTGAGAGAGCTGACACTGTTTGAACAGAGGCGCATGAG 1551  
QY 1584 AAGCTGTGCTATCTCAGAGCAAGAAAGTGTCTTGTGCGGCTCCGCTCTCAGTGT 1643  
DB 1552 AAGCTGTGCTATCTCAGAGCAAGAAAGTGTCTTGTGCGGCTCCGCTCTCAGTGT 1611  
QY 1644 GCAGCTGCGGCTGAGGCGAGTGAAGTATGCTGCTGAGAGAGTGTGCTGCGGCGG 1703  
DB 1612 GCAGCTGCGGCTGAGGCGAGTGAAGTATGCTGCTGAGAGAGTGTGCTGCGGCGG 1671  
QY 1704 GGAACCCCTATTTGCGCTGAGAGCGTCAACACAGCGGCTGTGAGGCGGCTGAGGCACTT 1763  
DB 1672 GGAACCCCTATTTGCGCTGAGAGCGTCAACACAGCGGCTGTGAGGCGGCTGAGGCACTT 1731  
QY 1764 TGGATCTTACTGATTCAGACATGTGATGACCTGGACACTTCAAGCAATTTGCAACTCG 1823  
DB 1732 TGGATCTTACTGATTCAGACATGTGATGACCTGGACACTTCAAGCAATTTGCAACTCG 1791  
QY 1824 TGGAGTGAAGTACAGTCAAGGCGCCACCTNCCCAAAATATCAGGTGTGAGGCGGCAAGA 1883  
DB 1792 TGGAGTGAAGTACAGTCAAGGCGCCACCTNCCCAAAATATCAGGTGTGAGGCGGCAAGA 1848  
QY 1884 CTTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1942  
DB 1849 CTTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1907

RESULT 11  
US-09-989-920-73  
Sequence 73, Application us/09989920  
Patent No. US20020172957A1  
GENERAL INFORMATION:  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Herve  
APPLICANT: Chen, Sei-Yu  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P  
FILE REFERENCE: DEX-0291  
CURRENT APPLICATION NUMBER: US/09/989,920  
PRIOR APPLICATION NUMBER: 2001-11-21  
PRIOR FILING DATE: 2000-11-22  
NUMBER OF SEQ ID NOS: 284  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 73  
LENGTH: 8095  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-989-920-73

Query Match 61.0%; Score 1394; DB 10; Length 8095;  
Best Local Similarity 97.6%; Pred. No. 0;  
Matches 1448; Conservative 0; Mismatches 30; Indels 6; Gaps 3;

QY 462 TATCCAGAAAGGAGAAACCAACCAAGACGAGTGTCTTAACCTTATCCGCTTCTGACGC 521  
DB 5887 TATCCAGAAACACTTCTTCCGCCAGACGAGTGTCTTAACCTTATCCGCTTCTGACGC 5946  
QY 522 CTCAAGGCTCCCACTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 581  
DB 5947 CTCAAGGCTCCCACTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 6006

QY 582 CGTCAACATGCTACCTTCACTTTGGAGCAGAGAGTTTGAAGATGGGAAGCGCAAGTG 641  
 DB 6007 CGTCAACATGCTACCTTCACTTTGGAGCAGAGAGTTTGAAGATGGGAAGCGCAAGTG 6066  
 QY 642 TCCCATATACCCAGCTAAGGGCCATGCGCCCTTTGGAGTGGTGGAGCTTACTCGGC 701  
 DB 6067 TCCCATATACCCAGCTAAGGGCCATGCGCCCTTTGGAGTGGTGGAGCTTACTCGGC 6126  
 QY 702 CACACTCAACAACTTCTGGGACCGGAACCCATTTCCCTGCTTAACATATGGGCCCCACCA 761  
 DB 6127 CACACTCAACAACTTCTGGGACCGGAACCCATTTCCCTGCTTAACATATGGGCCCCACCA 6186  
 QY 762 CTCCATGAAGACAGAGTACCTGGCCCTTTTGGCTCAACGAACTCACTTTGATGCTTCG 821  
 DB 6187 CTCCATGAAGACAGAGTACCTGGCCCTTTTGGCTCAACGAACTCACTTTGATGCTTCG 6246  
 QY 822 CTATGACTGAGATGTGGGACGCTTACGCGGGGACGACGACAGAGTCTACTTCTCT 881  
 DB 6247 CTATGACTGAGATGTGGGACGCTTACGCGGGGACGACGACAGAGTCTACTTCTCT 6306  
 QY 882 CAGGAGCGCGGCTGAGATCGGCTGATGCTCCAGAGAGTGTGGCTCTGTGGCCG 941  
 DB 6307 CAGGAGCGCGGCTGAGATCGGCTGATGCTCCAGAGAGTGTGGCTCTGTGGCCG 6366  
 QY 942 TGTCTGCAAGGGCCATATGGGGGGGCGACAGACCTGCGAGAGAAAGTGACCACTTCT 1001  
 DB 6367 TGTCTGCAAGGGCCATATGGGGGGGCGACAGACCTGCGAGAGAAAGTGACCACTTCT 6426  
 QY 1002 GAAGCGGGGCTGAGATGCTGCGCCGCAACTGGAGCTCTACTTCAACACAGCTGACAGC 1061  
 DB 6427 GAAGCGGGGCTGAGATGCTGCGCCGCAACTGGAGCTCTACTTCAACACAGCTGACAGC 6486  
 QY 1062 GATGACACACCTGACGACACCTCTGTGACACACACACACCTTTGGGGTTTTCAAGC 1121  
 DB 6487 GATGACACACCTGACGACACCTCTGTGACACACACACACCTTTGGGGTTTTCAAGC 6546  
 QY 1122 ACAGGGGGGTGACATGTAACCTGTGGCCATCTGTGATGATACCAAGTTGGAAAGATCCAGC 1181  
 DB 6547 ACAGGGGGGTGACATGTAACCTGTGTGGCCATCTGTGATGATACCAAGTTGGAAAGATCCAGC 6606  
 QY 1182 GGTGTTTGAAGGCGCCCTTAAAGAGTACCATGAGAGAACCCCGAAGTGGAGCCGCTACAC 1241  
 DB 6607 GGTGTTTGAAGGCGCCCTTAAAGAGTACCATGAGAGAACCCCGAAGTGGAGCCGCTACAC 6666  
 QY 1242 TGACCTGTACCCAGCCCTCGGCTGCTGTGATTAACAACATGGGCACTGGGCGCACGG 1301  
 DB 6667 TGACCTGTACCCAGCCCTCGGCTGCTGTGATTAACAACATGGGCACTGGGCGCACGG 6726  
 QY 1302 CTACACGAGCTCCCGGAGCTACCCGACAACTCTCAACTCTGTCAGAAAGACACCCGT 1361  
 DB 6727 CTACACGAGCTCCCGGAGCTACCCGACAACTCTCAACTCTGTCAGAAAGACACCCGT 6786  
 QY 1362 GATGAGAGAGAGTGGGCGCCCTGCTGAGAGCCGCGCTGCTGCTGTAAGAAAGGCGACCA 1421  
 DB 6787 GATGAGAGAGAGTGGGCGCCCTGCTGAGAGCCGCGCTGCTGCTGTAAGAAAGGCGACCA 6846  
 QY 1422 CTTCACCACTGCTGGCGGACCGGCTTACAGAGATTTATGAGGACCACTTATACGTCT 1481  
 DB 6847 CTTCACCACTGCTGGCGGACCGGCTTACAGAGATTTATGAGGACCACTTATACGTCT 6906  
 QY 1482 GTTCATTGGCAGAGAGAGGCTGCTCAAGGCTGTGAGGCTGGGGCCCTGGGTTCA 1541  
 DB 6907 GTTCATTGGCAGAGAGAGGCTGCTCAAGGCTGTGAGGCTGGGGCCCTGGGTTCA 6966  
 QY 1542 CCGTATGAGAGAGCTGAGCTGTTTGAACAGAGAGCCATGAGAAAGCTGTGCTATCTCA 1601  
 DB 6967 CCGTATGAGAGAGCTGAGCTGTTTGAACAGAGAGCCATGAGAAAGCTGTGCTATCTCA 7026  
 QY 1602 GAGCAAG--AAGCTGCTCTTTGGCGGCTCCGCTCTGAGCTGTGACAGCTGCGCCGTGAC 1658  
 DB 7027 GAGCAAGCTTAAAGCTGCTCTTTGGCGGCTCCGCTCTGAGCTGTGAGCTGCGCCGTGAC 7086  
 QY 1659 CGAGCTGAATAAGTATGCTCTCTGTGAGAGCTGTCTCTCCGCGCGGAGACCCCTATTTGGC 1718

DB 7087 CGACTGCAATGAAGTATCGCTCTCTGTGACAGACTGTGCTCGCGCGGAGACCCCTATTTGCGC 7146  
 QY 1719 CTGGAGCGCTCAACACACAGCGGCTGTGGCGCGTGGGTCGCACTTTGATCTTACTGAT 1778  
 DB 7147 CTGGAGCGCTCAACACACAGCGGCTGTGGCGCGTGGGTCGCACTTTGATCTTACTGAT 7206  
 QY 1779 CCAGCATGTGATGACTCGGACACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAGATACA 1838  
 DB 7207 CCAGCATGTGATGACTCGGACACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAG--AAA 7264  
 QY 1839 GTGAGCGCCACTNCCAAAACATCAAGCTGTGTGGCGGACAGACAGCTGTGCTGCTG 1898  
 DB 7265 GTGAGCGCCACT--CCAAAACATCAAGCTGTGTGGCGGACAGACAGCTGTGCTGCTG 7323  
 QY 1899 CCACCTCTCTCCAACTTGGCCCTGCGCGAGCTCCAAACCCCGAGG 1942  
 DB 7324 CCACCTCTCTCCAACTTGGCCCATGCGCGAGCTGTGAGG 7367

RESULT 12  
 US-09-759-130B-336  
 ; Sequence 336, Application US/09759130B  
 ; Publication No. US20030022279A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millennium Pharmaceuticals, Inc.  
 ; APPLICANT: McCarthy, Sean A  
 ; APPLICANT: Fraser, Christopher C  
 ; APPLICANT: Sharp, John D  
 ; APPLICANT: Barnes, Thomas S  
 ; APPLICANT: Kistel, Susan J  
 ; APPLICANT: Mackay, Charles R  
 ; APPLICANT: Myers, Paul S  
 ; APPLICANT: Leiby, Kevin R  
 ; APPLICANT: Wrighton, Nicolas  
 ; APPLICANT: Goodearl, Andrew  
 ; APPLICANT: Holzman, Douglas A  
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
 ; FILE REFERENCE: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER  
 ; FILE REFERENCE: MP100-5350WIM  
 ; CURRENT APPLICATION NUMBER: US/09/759,130B  
 ; CURRENT FILING DATE: 2002-09-16  
 ; PRIOR APPLICATION NUMBER: US 09/479,249  
 ; PRIOR FILING DATE: 2000-01-07  
 ; PRIOR APPLICATION NUMBER: US 09/559,497  
 ; PRIOR FILING DATE: 2000-04-27  
 ; PRIOR APPLICATION NUMBER: US 09/578,063  
 ; PRIOR FILING DATE: 2000-05-24  
 ; PRIOR APPLICATION NUMBER: US 09/333,159  
 ; PRIOR FILING DATE: 1999-06-14  
 ; PRIOR APPLICATION NUMBER: US 09/596,194  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: US 09/342,364  
 ; PRIOR FILING DATE: 1998-06-29  
 ; PRIOR APPLICATION NUMBER: US 09/608,452  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/393,996  
 ; PRIOR FILING DATE: 1999-09-10  
 ; PRIOR APPLICATION NUMBER: US 09/602,871  
 ; PRIOR FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: US 09/420,707  
 ; PRIOR FILING DATE: 1999-10-19  
 ; NUMBER OF SEQ ID NOS: 460  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 336  
 ; LENGTH: 3503  
 ; TYPE: DNA  
 ; ORGANISM: Mus sp.  
 ; US-09-759-130B-336  
 Query Match 59.2%; Score 1351; DB 11; Length 3503;  
 Best Local Similarity 84.7%; Pred. No. 0;







```

Db      1505  CCCCCGCGCTGCGGGAAGAGCTGAGAAAGGGGCGCAAGGCTACTGAGAGGACCT--TGG 1562
Oy      1550  AGGACCTGCAGCTGTTTACACAGAGACCCATGAGAAGCGTGTGCTATCTCAGACCAAGA 1609
Db      1563  TGTACCCCTGAGCTGCGCCCAAGAGAGCCACAGTCCCTCC--GGCCCTGTCTCTGA 1620
Oy      1610  AGCTCTCTTTTCCGCGCTCCGCTCTCAAGCTGTGACAGCTGCGCGTGGCCGACCTGATTA 1669
Db      1621  ACCAATATGAAAGAACTTTGGATCTCTGCGGTACTACTATTCAGATGCTGCTCCCTAAAGAT 1680
Oy      1670  AGTATGCTCTCTGTGACAGACTGTGCTCTGCGCGGAGCCCTATTTGCGGCTGTGAGCTGCA 1729
Db      1681  AGTACCTGGGCAATGCGCCGCTGCGACCGCGGTGGGGGCGCCCTTGGCCACCTCCAGGCAAT 1740
Oy      1730  ACACGAGCGCTGTGTGCGCGTGGGTGGCCACTTTGATCTTTACTGATCCAGATGTGA 1789
Db      1741  CCCAGGCGACCTGTGCGCTTCTCCAACTGAGGCTTACCGTGGGGGGTGGCGGCAACTCAAA 1800
Oy      1790  TGACCTGGACACTTCAGGCAATTTGCAACCTCCGTGGCAGTAAGTACAGTACGCGCCAC 1849
Db      1801  TGCCAAATGTTTACGTGCGCTTACAACTAGGAGGGAGGACCGGGGAGGCGCTGCGGACACC 1860
Oy      1850  TNCCTCAAAATCATCAGGTGTGTGGGCGACAGACCTGCTGCTGCGCTGCGCACCTCTCT 1909
Db      1861  CCTGCTGAGCTGCGCGGATGAACTAGAGAGCAAACTGCAGCAAGCCCAAGCA----- 1912
Oy      1910  CCMACTTGCGCTTGGCGCGCTCCAAACCCGAGAGATCATCACTATGAGGGGAACCCCGAC 1969
Db      1913  -----CTGCGCGACTCAACCCCGAGAGATCATCACTATGAGGGGAACCCCGAC 1961
Oy      1970  CGCGTGGCGGAGNACCGTGGGAGGTGTAGCTCTACTTTTGCACAGGACCAAGCTATCTC 2029
Db      1962  CGCGTGGCGGAGNACCGTGGGAGGTGTAGCTCTACTTTTGCACAGGACCAAGCTATCTC 2021
Oy      2030  AGGACATGAGCAGGAGCACTGCTGTCTGTGGGACAGATCTGCCAGACCCACCGCGC 2089
Db      2022  AGGACATGAGCAGGAGCACTGCTGTCTGTGGGACAGATCTGCCAGACCCACCGCGC 2081
Oy      2090  CATGAGACCTGCTGTCTGTCTGTGTGAGGACAGATCTGCCAGACCCACCGCGC 2147
Db      2082  CATGAGACCTGCTGTCTGTCTGTGTGAGGACAGATCTGCCAGACCCACCGCGC 2141
Oy      2148  AG-CTGCGAAGAGGATCTCTCTCTGTGTGATTCACAGACAGCGGGAGCCACCGCG 2206
Db      2142  AGCTGCGAAGAGGATCTCTCTCTGTGTGATTCACAGACAGCGGGAGCCACCGCG 2201
Oy      2207  CCAAAA-TTTTCAAGGACAGAG-TTMAAGATGTGTGTGATTTGACATGTGTG 2264
Db      2202  CCAAAACTTTTCAGAGGAGATTTTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2261
Oy      2265  TGTGTGTGTGTGTGTGTGTG 2284
Db      2262  TGTGTGTGTGTGTGTGTGTG 2281

```

## RESULT 15

```

US-10-189-123-33
; Sequence 33, Application US/10189123
; Publication No. US20030082586A1
; GENERAL INFORMATION:
; APPLICANT: KIRST, Susan J.
; APPLICANT: HOUTZMAN, Douglas A.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; APPLICANT: BARNES, Thomas S.
; TITLE OF INVENTION: ANTIBODIES HAVING DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: 10147-1103
; CURRENT APPLICATION NUMBER: US/10/189,123
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: US 09/342,364

```

```

; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 33
; LENGTH: 2811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-189-123-33

```

Query Match 52.3%; Score 1193.6; DB 14; Length 2811;

Best Local Similarity 74.6%; Pred. No. 0; Mismatches 423; Indels 161; Gaps 11;

```

Matches 1716; Conservative 0;

```

Oy	118	CGCACACGACGAGACAGCTGAGCTGAAAGCTCAGAGCCGGGCGCTGCGCATGAGCCACAC	177
Db	10	CGCGTGGGAGACAGCTGAGCTGAAAGCTCAGAGCCGGGCGCTGCGCATGAGCCACAC	69
Oy	178	TGGGCTGTGCTGCTGCTGCGACAGAGGCTGTGGGCTGTGGCATTTGGGCTAGTGTGG	237
Db	70	TGGGCTGTGCTGCTGCTGCGACAGAGGCTGTGGGCTGTGGCATTTGGGCTAGTGTGG	129
Oy	238	TGGAACCTTGTGCGCGTAAAGACAGTGTCTGTGGGACCTGGCCAGGTAGTACGGCG	297
Db	130	TGGAACTTGTGCTGCGCGTAAAGACAGTGTCTGTGGGACCTGGCCAGGTAGTACGGCG	189
Oy	238	TTCTCCAGACCGGATCCAGGACTTCTGTACACTGACCTGACGAGAGCCACTGGGCTT	357
Db	190	TTCTCCAGACCGGATCCAGGACTTCTGTACACTGACCTGACGAGAGCCACTGGGCTT	249
Oy	358	CTGTACGTGGGCGCGGAGAGGCGCTGTGCTTGTGCTTACAGATGAGAGCCCTGAGCTGCA	417
Db	250	CTGTACGTGGGCGCGGAGAGGCGCTGTGCTTGTGCTTACAGATGAGAGCCCTGAGCTGCA	309
Oy	418	GGAGGATCTCTGTGGAGGCGCCCGTGGAGAAAGACTGATGTATCCAGAAAGGAAG	477
Db	310	GGAGGATCTCTGTGGAGGCGCCCGTGGAGAAAGACTGATGTATCCAGAAAGGAAG	369
Oy	478	AACAACAGACCGAGTCTTCAATTCATCGCTTCTGTGACAGCCCTTACAAATGCTCCAC	537
Db	370	AACAACAGACCGAGTCTTCAATTCATCGCTTCTGTGACAGCCCTTACAAATGCTCCAC	429
Oy	538	CTGTACGTGTGGGACCTGACAGGCTTCCAGCCCAAGTGCACCTAGTCTGTC-----	585
Db	430	CTGTACGTGTGGGACCTGACAGGCTTCCAGCCCAAGTGCACCTAGTCTGTC-----	489
Oy	586	-----	585
Db	490	CTCTACCTCGGTGTCCCGAGCCCGCCCTCTCACCCCTTCTGTGACTGTGATGT	549
Oy	586	-----	585
Db	550	GGCCCAAGAGCCCTGCGCTTAAGCATCTCTCATACCTCTCTGTCTCTTAAGACA	609
Oy	590	TGCTACCTTCACTTTGAGAGCATGAGAGTTGAAGATGGAAGGCAAGTGTCCCTATG	649
Db	610	TGCTACCTTCACTTTGAGAGCATGAGAGTTGAAGATGGAAGGCAAGTGTCCCTATG	669
Oy	650	ACCGAGCTAAGGCGCATGCTGCGCTTCTGTGATGTGTGAGTGTACTGTGCGCACATCA	709
Db	670	ACCGAGCTAAGGCGCATGCTGCGCTTCTGTGATGTGTGAGTGTACTGTGCGCACATCA	729
Oy	710	ACAACTTCTGTGGGACAGGAACCATTTATCTGCTGAACATGAGGCGCCACCACTCATGA	769
Db	730	ACAACTTCTGTGGGACAGGAACCATTTATCTGCTGAACATGAGGCGCCACCACTCATGA	789
Oy	770	AGCAGAGTACTGCGCTTTTGGCTCAAGCAAGCTCACTTGTAGGCTGTGCTTGTGAC	829
Db	790	AGCAGAGTACTGCGCTTTTGGCTCAAGCAAGCTCACTTGTAGGCTGTGCTTGTGAC	849
Oy	830	CTGAGAGTGTGGGACCTTACAGGGGAGAGCAAGAGTGTACTTCTTACAGGAGC	889
Db	850	CTGAGAGTGTGGGACCTTACAGGGGAGAGCAAGAGTGTACTTCTTACAGGAGC	909

QY	890	GGGAGAGTGGATCCGACGCGTATATCCGAGCAGAGTGGTGGCGTCTGGTGGCCCGCGTCTGCA	949
Db	910	GGGCAATGGATCCGACTGCTATATCCGAGCAGAGTGGTGGCGTCTGGTGGCCCGCGTCTGCA	969
QY	950	AGGCGATATGGGGGGGGGCGACAGGACCCCTGCGAGAGAGTGGACCCAGCTTCTCGAAGCGC	1009
Db	970	AGGCGATATGGGGGGGGGCGACAGGACCCCTGCGAGAGAGTGGACCCAGCTTCTCGAAGCGC	1029
QY	1010	GGCTGGCATGCTCTGCCCCGACCTGGCAGCTCTACTTCAMACAGCTGCAGGCGATGCACA	1069
Db	1030	GGCTGGCATGCTCTGCCCCGACCTGGCAGCTCTACTTCAMACAGCTGCAGGCGATGCACA	1089
QY	1070	CCCGGAGGAGACCTCCGCGGACAAACACACTCTTTGGGGGTTTTCAACAGCAAGTGGG	1129
Db	1090	CCCGGAGGAGACCTCCGCGGACAAACACACTCTTTGGGGGTTTTCAACAGCAAGTGGG	1149
QY	1130	GTGACATGTACTCTCGGCGCATCTGTGAGTACAGATTGGGAAGATCAGAGGGGTGTTTG	1189
Db	1150	GTGACATGTACTCTCGGCGCATCTGTGAGTACAGATTGGGAAGATCAGAGGGGTGTTTG	1209
QY	1190	AGGGCCCCCTATAGGAGTACCATTAGGAAAGCCAGAAAGTGGAGCCGTACACTGACCCTG	1249
Db	1210	AGGGCCCCCTATAGGAGTACCATTAGGAAAGCCAGAAAGTGGAGCCGTACACTGACCCTG	1269
QY	1250	TACCCAGCCCTCGGCGCTGCTGTCGTGCTATTAACTGGCATCGGCGCCACGGCTACACA	1309
Db	1270	TACCCAGCCCTCGGCGCTGCTGTCGTGCTATTAACTGGCATCGGCGCCACGGCTACACA	1328
QY	1310	GCTCCCTGGAGCTACCCGACAAACATCCCTCAACCTTCGTCAACAAACACCCCGTGATGGAG	1369
Db	1329	TTTCAAGAGGACAGAGGGGCGCGGCTGGCTGCTGAAGCTTACTTGTGGCTCTCGTGGAG	1388
QY	1370	AGCAGTGGGGGCTCGGTGAGAGCCGCCCTGCTGTTGAAAGAGGAGCCACCACTTCAACC	1429
Db	1389	GCCGCTGGGTGACCTTGGAGAGCCCGGGGCCCCCTGGAAACCTGAGGGGTGGTGGGTGG	1448
QY	1430	ACCTGCTGGCCGACCGGTTACAGAGCTTGAATGAGCCACTATACAGTCTGTTTCATG	1489
Db	1449	CGGTGGTGGCCCTGGGGGCTGTGTGCC---TGGTGTGCTGCTGCTGGTGTGCTGCTATG	1504
QY	1490	GCACAGAGAGAGGGGTGGCTGCTCAAAAGGCTGTGAGCCTGGGGCCCTGGGGTTACCTGATG	1549
Db	1505	CGCCGGCGGCTGCGGGGAAAGCTGGAGAAAGGGGCCAAAGCTTACTGAGAGCACT--TGG	1562
QY	1550	AGGAGCTGACACTTTGTGACGAGAGGCCATAGAAAGCTGCTATCTACAGACAGA	1609
Db	1563	TGTACCCCTGAGAGTGGCCAAAGAGGCCAACAGTCCCCCTTCC--GGCCCTGTCTGA	1620
QY	1610	AGCTGCTCTTTGCCGGCTCCCGCTCTACGCTGTGTGACAGCTGCCCGTGGCCGACTGCATAA	1669
Db	1621	ACCAGATGAGAACTTTGGGATCCTGTGCTGTACTACTATTCAGATGGCTCCCTTAAGAT	1680
QY	1670	AGTATCGTCTCTGTGACAGACTGTGCTCTCGCCCGGGAGCCCTATTTGGCCCTGGAGCGTCA	1729
Db	1681	AGTATCGTGGGATGCGCCGCTGCCAGAGCCCGGGGGGGGCCCTTCGCGACCTCCAGGAT	1740
QY	1730	ACACAGCCGCTGTGTGGCCGTGGGTGGCCATTGTGATCTTACTATGCAGCATGGA	1789
Db	1741	CCGAGGCGACGCTCTGCTCTTCCAACTCGGCTTACCTGGGGGGTGGCGGAATCTCAA	1800
QY	1790	TGACCTCGGACACTTACAGGCAATTTGCAACCTCCGTGGCATGAATACATGTCAGGCCAC	1849
Db	1801	TGCCAATGTGTACGTGCGCTTACAACTAGAGGGGAGGACCGGGGAGGGCTCGGGACACC	1860
QY	1850	TNCCCAAAATCATCAGGTGTGGTGGCGGACAGACCTGTGTGCTGCGCTCCGCACTCTCT	1909
Db	1861	CTGTGCTGAGTCTCGGGATGAACTGACAGCGAAACTGCAGCAACGCCAGCA-----	1912
QY	1910	CCAACTTGGCCCTGCCGACTCAACACCCGAGAGTCACTAGTTAGAGGGAACCCCAAC	1969
Db	1913	-----CTGGCGGACTCAACACCCCGAGAGTATATGATAGAGGGAACCCCAAC	1961
QY	1970	CGCGTGGCGGAGNAGCTGGGAGGCTGTAAGTCTTACTTTTGCACAGCACCAGCTATCTC	2029

Search completed: August 19, 2003, 17:23:14  
Job time : 427 secs

Db	1962	CGCGTCGGCGGAAAGCGTGGAGGTGTAGCTCTACTATTTCACAGGCAACGATACCTC	2021
QY	2030	AGGGACATATGGCAGCGGACACCTGCTCTGTCTTGGGACAGATACTGCCCCACACCCAGC	2089
Db	2022	AGGGACATAGGCAGCGGACACCTGCTCTGTCTTGGGACAGATACTGCCACACCCAGCGC	2081
QY	2090	CATAGAGACCGTCCTGCTGCACGACGGGCATCG-CACTTGGTGGG-TCACAGAGGCAC	2147
Db	2082	CATAGAGACCGTCCTGCTGCACGACGGGCATCGCACTTGGTGGGCTCACAGAGGCAC	2141
QY	2148	AG-CTCCGAGAGGCATCTTCCCTCCTCTGTGTGAATCACAGACAGCGCGGACCCACGCG	2206
Db	2142	AGCCTCGAGAGGCGATCTTCTCCTCTCTCTGTGAATCACAGACAGCGGAGCCACGCG	2201
QY	2207	CCAAAA-TTTTCAAGGCAGAG-TTMAAGATGTGTTTGTATTTGACATGTGTTTG	2264
Db	2202	CCAAAACTTTTCAAGGCAGAGATTTCAGATGTGTTTGTCTGTATTTGCACATGTCTT	2261
QY	2265	TGTGTGTGTATGTGTGTG 2284	
Db	2262	TGTGTGTGTGTATGTGTG 2281	

**THIS PAGE BLANK (USPTO)**